

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 17:12:45 ; Search time 354.405 Seconds  
(without alignments)  
3511.351 Million cell updates/sec

Title: US-10-019-783-1

Perfect score: 2390

Sequence: 1 MVHQSNHGGERAAAAANGKS.....LKVKSFCQRNKKNSINGC 461

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DB=xlh

-O=/cpn2.1/USPTO/spec/US10019783/runat\_07112003\_161642\_19826/app\_query.fasta\_1.1358

-DB=N\_Geneseq 15Jun03 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCPU=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000

-USER=US10019783@CGN 1.1.458 @runat\_07112003\_161642\_19826 -NCPU=6 -ICPU=3

-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq 19Jun03:\*\*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*\*

2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*\*

3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*\*

4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*\*

5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*\*

6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*\*

7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*\*

8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*\*

9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*\*

10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*\*

11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*\*

12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*\*

13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*\*

14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*\*

15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*\*

16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*\*

17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*\*

18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*\*

19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*\*

20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*\*

21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*\*

22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*\*

23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*\*

24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*\*

25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	1660	19	Nicotianamine amin
2	2142.5	89.6	1910	19	Nicotianamine amin
3	1988	83.2	10966	22	Hordeum vulgare L.
4	1124	47.0	1389	24	Arabidopsis thalia
5	1023.5	42.8	1338	24	Arabidopsis thalia
6	1023.5	42.8	1533	21	Arabidopsis thalia
7	1017.5	42.6	1535	21	Arabidopsis thalia
8	872	36.5	1481	21	Arabidopsis thalia
9	865.5	36.2	1484	21	Arabidopsis thalia
10	793	33.2	1584	21	Arabidopsis thalia
11	792.5	33.2	1101	21	Arabidopsis thalia
12	792.5	33.2	1182	21	Arabidopsis thalia
13	788.5	33.0	1184	21	Arabidopsis thalia
c 14	649	27.2	858	24	Arabidopsis thalia
15	647	27.1	1377	25	Rat tyrosine amino
16	646.5	27.1	1844	23	Drosophila melanog
17	644	26.9	2754	24	Gene #3723 used to
18	644	26.9	2782	22	Human liver cell s
c 19	644	26.9	4412	22	Human heart cell s
20	564.5	23.6	5492	23	Drosophila melanog
21	492	20.6	1024	25	Toxicologically re
22	361.5	15.1	1185	18	DNA encoding Aquif
23	361.5	15.1	1185	25	Aquifex aspartate
c 24	361	15.1	4403765	22	Mycobacterium tube
c 25	361	15.1	4411529	22	Mycobacterium tube
26	358	15.0	48275	21	N. meningitidis pa
c 27	358	15.0	349980	21	Neisseria meningit
c 28	358	15.0	349980	21	Neisseria meningit
c 29	358	15.0	837096	21	N. meningitidis pa
30	355	14.9	1434	22	Corynebacterium gl
31	355	14.9	1434	22	Corynebacterium gl
32	355	14.9	309400	22	C glutamicum codin
33	355	14.9	349980	22	C glutamicum codin
34	354	14.8	1212	25	N. gonorrhoeae nuc
35	354	14.8	1257	22	C glutamicum codin
c 36	346	14.5	349980	22	Pyrococcus abyssi
c 37	338.5	14.2	48328	22	Genomic fragment #
c 38	331	13.8	370	25	Human GDP-mannose
c 39	329.5	13.8	818	22	Human breast cance
c 40	329.5	13.8	2365589	24	Genomic sequence o
c 41	327	13.7	1163020	24	Listeria innocua c
c 42	327	13.7	3011208	24	Listeria innocua D
c 43	323	13.5	10863	22	DNA encoding novel
44	323	13.5	349980	24	Bifidobacterium lo
c 45	321.5	13.5	2944528	24	Listeria monocytog

ALIGNMENTS

RESULT 1

AAV48147

ID AAV48147 standard; CDNA to mRNA; 1660 BP.

XX

AC AAV48147;

XX

DT 27-OCT-1998 (first entry)

XX

DE Nicotianamine aminotransferase 49564.15 molecular weight protein, gene.

XX ds; nicotianamine aminotransferase; plant; iron absorption;

KW iron deficiency chlorosis.

XX

OS Gramineae sp.

XX

XX Key

PH Location/Qualifiers

FT 62..1447

```
FT / *tag= a
FT /product= "Nicotianamine aminotransferase"
```

```

FT      /*tag= a
FT      /product= "Nicotianamine aminotransferase"
XX
XX      EP860499-A2.
XX
XX      26-AUG-1998.
XX
XX      19-FEB-1998; 98EP-0102891.
XX
XX      21-FEB-1997; 97JP-0037499.
XX
XX      (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX      Mori S, Nakanishi H, Takahashi M;
XX
XX      WPI; 1998-439341/38.
XX      P-PADB; AAW61642.
XX
XX      New nicotianamine aminotransferase protein and DNA - useful for
XX      enhancing iron absorption of plant cells
XX
XX      Claim 4; Page 12-13; 17pp; English.
XX
XX      The nicotianamine aminotransferase can be used in a plasmid to transform
XX      plant cells to produce cells with enhanced iron absorption, and it is
XX      implied (though not stated) that plants with improved resistance to iron
XX      deficiency chlorosis in calcareous soils can be regenerated from the
XX      transformed cells. The gene fragment can be used to detect, amplify
XX      and/or isolate nicotianamine aminotransferase genes.
XX
XX      Sequence 1660 BP; 423 A; 442 C; 430 G; 365 T; 0 other;
SQ

Alignment Scores:
Pred. No.:      1.33e-205      Length:      1660
Score:      2390.00      Matches:      461
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      19      Gaps:      0

US-10-019-783-1 (1-461) x AAV48147 (1-1660)

```

DE Nicotianamine aminotransferase 58148.62 molecular weight protein, gene.  
 XX  
 KW ds; nicotianamine aminotransferase; plant; iron absorption;  
 KW iron deficiency chlorosis.  
 XX  
 OS Gramineae sp.  
 XX  
 EH Key Location/Qualifiers  
 FT 76..1731  
 FT /\*tag= a  
 FT /product= "Nicotianamine transferase"  
 XX  
 XX EP860499-A2.  
 XX  
 XX 26-AUG-1998.  
 XX  
 XX 19-FEB-1998; 98EP-0102891.  
 XX  
 XX 21-FEB-1997; 97JP-0037499.  
 XX  
 XX (SUMO) SUMITOMO CHEM CO LTD.  
 XX  
 PI Mori S, Nakanishi H, Takahashi M;  
 XX WPI; 1998-439341/38.  
 DR P-PSDB; AAW61643.  
 DR  
 XX New nicotianamine aminotransferase protein and DNA - useful for  
 PT enhancing iron absorption of plant cells  
 PT  
 XX Claim 4; Page 14-15; 17pp; English.  
 XX  
 CC The nicotianamine aminotransferase can be used in a plasmid to transform  
 CC plant cells to produce cells with enhanced iron absorption, and it is  
 CC implied (though not stated) that plants with improved resistance to iron  
 CC deficiency chlorosis in calcareous soils can be regenerated from the  
 CC transformed cells. The gene fragment can be used to detect, amplify  
 CC and/or isolate nicotianamine aminotransferase genes.  
 XX  
 SQ Sequence 1910 BP; 462 A; 534 C; 546 G; 368 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 3,228-183 Length: 1910  
 Score: 2142.50 Matches: 415  
 Percent Similarity: 93.55% Conservative: 20  
 Best Local Similarity: 89.25% Mismatches: 21  
 Query Match: 89.64% Indels: 9  
 DB: 19 Gaps: 3  
 US-10-019-783-1 (1-461) x AAV48148 (1-1910)  
 QY 4 GlnSerAsnGlyHisGlyGluAlaAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHis 23  
 Db 340 GAGGACAAACGGCCATGCTGAG-----GCCGCCGACGGCAACGGCGGACGAGCAT 393  
 QY 24 Ala-----AlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaAlaAlaAla 38  
 Db 394 GCGGAGGACTCCGGCGGCAACGGCGAGCAACGGGCATGCGCGCGCGCGCGGACGAGAG 453  
 QY 39 -----AlaValGluTrpAsnPheAlaAlaArgGlyLysAspGlyIleLeuAlaThrThrGly 56  
 Db 454 GAGGAGGCGGTGGAGTGGAAATTCGCGGTGCCAAGGACGGCGTCTCGCGCGGACGCGG 513  
 QY 57 AlaLysAsnSerIleArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGly 76  
 Db 514 GCGAATCATGAGCATCCGGCGGATAGCGTACAGATCAGCGGAGCGGTGCGAGGAAAGGGG 573  
 QY 77 ProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThr 96  
 Db 574 CCGCGGCGCGTCTCGCGCTGCGCCACCGGACCCCGTCCGTTCCCGGCGCTTCGCGACG 633  
 QY 97 AlaValGluAlaGluAspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 116  
 Db 1714 TCGAGCGATGATTGC 1728

Db 634 GCGTCTGAGCGCGGAGGACGCGCTGCGCGCGCGTGGCGACCGCGCAGTTCAACTGCTAC 693  
 QY 117 AlaAlaGlyValGlyLeuProAlaAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGly 136  
 Db 694 CCGCGCGCGTGGCGCTCCCGCGCGGACGAAAGCGCGTGGCGAGCACCTGTCCGAGGC 753  
 QY 137 ValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIle 156  
 Db 754 GTGCGGTACATGCTATCGCGCGAGCGTCTTCTCACCGCGCGGAGCCCGAGGCGATC 813  
 QY 157 GluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgPro 176  
 Db 814 GAGTCAATAATCCCGGTGCTGGCCAGACCGCGCGGCCCAACATTCTGCTCCCCAGGCCA 873  
 QY 177 GlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAsp 196  
 Db 874 GGCTACCCAAACTACGAGGCGCGCGCGGTTCACAGGCTGGAGGTCCGCGCATTTTCGAC 933  
 QY 197 LeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLys 216  
 Db 934 CTCATCCCGACAAAGGGTGGGAGATCGACATCGACTCTGGAATCCATCGCGGCAAG 993  
 QY 217 AsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyr 236  
 Db 994 AACACCACCGCCATGCTCATATAAAACCCCAACACCGCTGCGCGACGCTTTACTCTCTAC 1053  
 QY 237 AspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAsp 256  
 Db 1054 GACCATCTGCCAAGTCCGAGGTGGCGAAGGCTCGAATATTTGGTATGCTGAC 1113  
 QY 257 GluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly 276  
 Db 1114 GAGTATACGGCAAGTGGTCTTGGCGAGCGCCCGTTCATCCCAATGGGAGTGTGGG 1173  
 QY 277 HisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrp 296  
 Db 1174 CATCATCCCTGTGCTGCTCCATAGGTCTCTCCAAAGTCATGGATAGTCCCTGGATGG 1233  
 QY 297 ArgLeuGlyTrpValAlaValTyrAspProThrLysIleLeuGluLysThrLysIleSer 316  
 Db 1234 CGCTTGGATGGGTAGCGGTGTACGCCGCCAAGAAAGATCTTCAGGAAACTAGATCTCT 1293  
 QY 317 ThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAla 336  
 Db 1294 ACATCAATTACGAATTACCTCAATGTCTCGACAGACCCAGCAACCTTCATTTCAGGAGCT 1353  
 QY 337 LeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeu 356  
 Db 1354 CTTCCTCAGATTCTTGAGAACACAAAGGAGATTCTTTAAGCGGATTATGGTCTGCTA 1413  
 QY 357 LysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysPro 376  
 Db 1414 AAGGAATCATCAGAGATATGCTACAAACAAATAAAGGAAACAAATACATTACATGTCTCT 1473  
 QY 377 HisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGlu 396  
 Db 1474 CACAAGCCAGAGGATCAATGTTTTCATGTTGTAAGTGAAGTGAAGTCACTTTTGGAGGA 1533  
 QY 397 IleHisAspAspIleAspPheCysCysLysLeuAlaLysGluLeuSerValIleLeuCys 416  
 Db 1534 ATAGACGATGACATTGATTTTCTGCAAGCTCGCAAGAGCAATCAGTAATCTTATGC 1593  
 QY 417 ProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSer 436  
 Db 1594 CCAGGAGGTGTTCTTGAATGGCAAACTGGGTCCGATTACTTTTCTGTGTTCATCT 1653  
 QY 437 SerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysAsn 456  
 Db 1654 TCTCTTCAGATGGTCTCGAGAGATCAATCAATCTCTGTCAAGGACCAAGAGAGATAT 1713  
 QY 457 SerIleAsnGlyCys 461  
 Db 1714 TCGAGCGATGATTGC 1728





Db 7715 TCATCTCTCATATGCTACTGCTGGGATTAGTATTTTGTCTAAATTTGTACTGCCCTTG 7774  
Qy 316 -----SerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThr 331  
Db 7775 TTATTCAGATCTCTACGCTCTATTACGAATTACCTTAATGTCTCAACGGACCCAGCAACC 7834  
Qy 332 PheValGln----- 334  
Db 7835 TTCGTTTCAGTTAGTCCTTTGGTCTCTGCGCTATTTTGTCTCATGCTCCCTGTGTGCAATGC 7894  
Qy 334 ----- 334  
Db 7895 AAATGACCGGCTTCAAGTTAGTATATAGAGTTTGTGTTAAGTGTGAATGTCGAAGTCCAA 7954  
Qy 334 ----- 334  
Db 7955 CATGATGGAAGAAAGATACATCTATTTTGTAGTCATTCGCCCTTTGTGTTTGTGATTCGATA 8014  
Qy 334 ----- 334  
Db 8015 AAATAAATAAACACAAAGCCAGAACCACTATTGTAATAGAACTATTTTCTTAGAAATA 8074  
Qy 334 ----- 334  
Db 8075 TACATTGATTTTGTAGCATGCGCATATCTTTTCGATCAAGTATGCAATATATTAAACTT 8134  
Qy 334 ----- 334  
Db 8135 GCATTGTACTAGAGTATACCATGTTGTGTTAAGAAATTTCTTACCTACACACCTTGTCTC 8194  
Qy 334 ----- 334  
Db 8195 GCATCTTCATATTTGATATCCTTGACATATTGTTCTCTTATGATTCACACAACTTAAT 8254  
Qy 335 -----GluAlaLeuProLysIleLeuGluAsnThr 344  
Db 8255 TATGATTTTGTGCTATCAAAATGTTAGGAAGCTCTCTCTAAATTTCTTGAGAACACA 8314  
Qy 345 LysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 364  
Db 8315 AAAGCAGATTCTTTAAGAGGATTATTGCTCTACTAAAGGAATCATCAGACATATGTTAT 8374  
Qy 365 ArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGluLysSerMetPhe 384  
Db 8375 AGGGAATAAGGAAACAAATATATTACGTGCTCTCAAGCCAGAGGATCGATGTTT 8434  
Qy 385 ValMet----- 386  
Db 8435 GTAATGGTAAGCTAAGCATAGACTTACTTTTAAAGTTAACTGGGATCTCAGTGCATCC 8494  
Qy 386 ----- 386  
Db 8495 AACAAACAATCAAAATCAAAATATATTATTTGCTATGGATCTTTTGAAGATGCATG 8554  
Qy 386 ----- 386  
Db 8555 CATTTGAAGAAATATGAAGAGATTGAATATTTTAGGACTAATCTCTCGATATCATTT 8614  
Qy 386 ----- 386  
Db 8615 TGTCCATTTTGTATTACTGTAAATTTGGTAACACTCAAAATCATATTACAAAAGTTT 8674  
Qy 386 ----- 386  
Db 8675 CCTCCATTTTAGTAAGATTGACTTCTTCTATATAACCATGTTAATTAATTCATGTA 8734  
Qy 387 ---ValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCys 405  
Db 8735 CAGGTCAAACTAACTTACATCTTTTGAGAGGATCCATGACGACATAAAATTTTGTGTC 8794  
Qy 406 LysLeuAlaLysGluSerValIleLeuCysPro----- 417

Db 8795 AAGTCGCAAGGAAGATCTGTAAATTTATGTCCAGTAGGAATGTATATGCGCAATTTT 8854  
Qy 417 ----- 417  
Db 8855 AAAGGAAACTATATGTAATAATAATATCTTCTTTATATAACAATACTTCTCCCAT 8914  
Qy 417 ----- 417  
Db 8915 CTTAAATAAATGTCTTACACTTAGCACAAATTTTATATAGATCTAGTACAAAAGTTGAAA 8974  
Qy 417 ----- 417  
Db 8975 CAGTTATTTGGACAGAGGAGTGTATATATTTGTGTGAGAACATAAAGTTATGTTGA 9034  
Qy 417 ----- 417  
Db 9035 CTGATATATGCTTCTTAAATGTGAAACATGTTCTCTTATGTTTTTTGATTGTATACGAAG 9094  
Qy 417 ----- 417  
Db 9095 TTCTTATCAGTTTCCGAGATGACTACACATAAATGATTACCATATCATTTGTCAGAAAATG 9154  
Qy 417 ----- 417  
Db 9155 TATTACCATAGATAATTTCTTTCTTTTATGCAAGAGCTAGCATGGCATGTACTTTTC 9214  
Qy 417 ----- 417  
Db 9215 CTTGTACCTATGTGCTTTTCTTCTCGTTACATGTTGTGCTTCTCACAAAAATAA 9274  
Qy 417 ----- 417  
Db 9275 ATACCAAGCACATGTTCCAAATGATTATTAATAATTTGAGGTGTTTTTCAACCAACTTA 9334  
Qy 417 ----- 417  
Db 9335 TATATCTCATAGTCTTAAAAAACGCTATATATGTTAACTCTCTAACAAAACTTATATA 9394  
Qy 418 -----GlySerValLeuGlyWetGluAsnTrpValArgIleThrPhe 431  
Db 9395 TGTCTCTCTCTAATACAGGAGTGTCTTGTGAATGAAATTTGGTCCGTATTACTTTT 9454  
Qy 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArg 451  
Db 9455 GCTGTGTTCCATCTTCTCTTCAAGATGGACTCGAAAGGTCAAATCATTTCTGTCAAAGG 9514  
Qy 452 AsnLysLysLysAsnSerIleAsnGlyCys 461  
Db 9515 AACAAAGAGAGATTCTATTAATGTTGT 9544  
RESULT 4  
ABZ14179  
ID ABZ14179 standard; DNA; 1389 BP.  
XX AC ABZ14179;  
XX 21-JAN-2003 (first entry)  
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1984.  
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX OS Arabidopsis thaliana.  
XX PN WO200216655-A2.  
XX PD 28-FEB-2002.  
XX PF 24-AUG-2001; 2001WO-US26685.  
XX PR 24-AUG-2000; 2000US-227866P.  
XX PR 26-JAN-2001; 2001US-264647P.  
XX PR 22-JUN-2001; 2001US-300111P.

(SRI ) SCRIPTS RES. INST.  
(SYGN ) SYNGENTA PARTICIPATIONS AG.  
Harper JF, Kreps J, Wang X, Zhu T;  
WPI; 2002-304127/34..  
Identifying a stress condition to which a plant cell has been exposed  
and producing plants with increased tolerance to these abiotic stresses  
Claim 144; SEQ ID NO 1984; 577pp + Sequence Listing; English.  
The invention relates to identifying a stress condition to which a plant  
cell has been exposed, comprising:  
(a) contacting nucleic acid representative of expressed polynucleotides  
in the plant cell with an array or probes representative of the plant  
cell genome; and  
(b) detecting a profile of expressed polynucleotides in the plant cell  
characteristic of a stress response. The method is useful in the  
production of transgenic plants, cells and seeds and in producing plants  
with increased tolerance to abiotic stress. The present sequence is that  
of an Arabidopsis thaliana stress regulated gene (ABZ1198-ASZ11574) used  
in methods of the invention.  
Note: The sequence data for this patent is not represented in the printed  
specification but is based on sequence information supplied to Derwent by  
the European Patent Office.

Alignment Scores:	
Pred. No.:	1.23e-91
Score:	1124.00
Percent Similarity:	66.52%
Best Local Similarity:	47.10%
Query Match:	49.03%
DB:	24
	24
	Gaps:
	4
	12
	138
	87
	211
	1389
	Length:
	Matches:
	Conservative:
	Mismatches:
	Indels:
	Gaps:
US-10-019-783-1 (1-461)	X ABZ14179 (1-1389)

XX WO200216655-A2.  
 XX 28-FEB-2002.  
 XX 24-AUG-2001; 2001WO-US26685.  
 XX 24-AUG-2000; 2000US-227866P.  
 XX 26-JAN-2001; 2001US-264547P.  
 XX 22-JUN-2001; 2001US-300111P.  
 XX (SCRI) SCRIPPS RES INST.  
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.  
 XX Harper JF, Kreps J, Wang X, Zhu T;  
 XX WPI; 2002-304127/34.  
 XX Identifying a stress condition to which a plant cell has been exposed  
 XX and producing plants with increased tolerance to these abiotic stresses  
 XX .  
 XX Claim 144; SEQ ID NO 1869; 577pp + Sequence Listing; English.  
 XX The invention relates to identifying a stress condition to which a plant  
 XX cell has been exposed, comprising:  
 XX (a) contacting nucleic acid representative of expressed polynucleotides  
 XX in the plant cell with an array or probes representative of the plant  
 XX cell genome; and  
 XX (b) detecting a profile of expressed polynucleotides in the plant cell  
 XX characteristic of a stress response. The method is useful in the  
 XX production of transgenic plants, cells and seeds and in producing plants  
 XX with increased tolerance to abiotic stress. The present sequence is that  
 XX of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used  
 XX in methods of the invention.  
 XX Note: The sequence data for this patent is not represented in the printed  
 XX specification but is based on sequence information supplied to Derwent by  
 XX the European Patent Office.  
 XX SQ Sequence 1338 BP; 375 A; 307 C; 326 G; 330 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.33e-82 Length: 1338  
 Score: 1023.50 Matches: 200  
 Percent Similarity: 66.51% Conservative: 84  
 Best Local Similarity: 46.84% Mismatches: 138  
 Query Match: 42.82% Indels: 5  
 DB: 24 Gaps: 2

US-10-019-783-1 (1-461) x ABZ14064 (1-1338)

QY 26 AlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaValGluTrpAsnPhaAla 45  
 DB 4 GCGAGCAACGAGTTACCACTGTAACGCAACGCCAATGTT-----TGGCGGTTC--- 54  
 QY 46 ArgGlyLysAspGlyLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArg 65  
 DB 55 -----AAAGGAAACGGTGCACACGAGTATCGCGCGGTGACGTTGACAAAGCTTGCT 108  
 QY 66 TyrLysIleSerAlaSerValGluSerGlyProArgProValLeuProLeuAlaHis 85  
 DB 109 TTTCGGATGTTAAAACTGCACCATGACAGTGAAGACCATTTTGTCCCACTCC 168  
 QY 86 GlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAla 105  
 DB 169 GCGAGCCCTCCGCCCATTCACACTTCAGGACTTGGCCGGAAGCGGAGGCGGTGTC 228  
 QY 106 AlaAlaLeuArgThrGlyGlnPheAsnGlyCysTyrAlaAlaGlyValGlyLeuProAlaAla 125  
 DB 229 GCACGTGCACCTCCGGCATCGCTAACTCTTACCGACCCCGCCCTGGAGTTTCAAGGCT 288  
 QY 126 ArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAsp 145

DB 289 AGAAGGCGGTGGCTGAATATTTAAACGGAGAACTTCCGACGAAGCTGAAGCGCGAGGAT 348  
 QY ValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProValLeuAlaGln 165  
 DB 349 GTGTATATCACCGAGAGTGTAAACAGCATAGATCGTGATAGATCTCTTTCGCGGA 408  
 QY ThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAla 185  
 DB 409 AATCCATCCGCCAACATTTCTACTTCCAAGCGCGGATATCTCTCACTACGATGCTGCT 468  
 QY 186 AlaPheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIle 205  
 DB 469 GTCTATAGCGCTCGAGATCCAAATAGATCTTCTCCCGAGAGTATTCGGAATC 528  
 QY 206 AspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsn 225  
 DB 529 AATCTCGATGGCTCGAGCGGCTCGGATGAGATACCGTCGAATGTTATCATCAAC 588  
 QY 226 ProAsnAspProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluVal 245  
 DB 589 CCAACAATCCATGTGGAACGCTACCTACGACCATCTCAACAGGTCCGCGGAGATG 648  
 QY 246 AlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGly 265  
 DB 649 GCTAGAAACTCGTATATGATAATATCCGACGAAGTATATGATCATGTTGTATATGGA 708  
 QY 266 SerAlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGly 285  
 DB 709 GACAAACCTTTATATCCCATGGGAAAGTTTGCATCAATAGCTCCGCTGATCAAGCTCGGA 768  
 QY 286 SerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAsp 305  
 DB 769 TCCATATCCAAAGGATGGGTCAACCGCTCGAGAGTTGGCTGGATGCCATGAACGAT 828  
 QY 306 ProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 325  
 DB 829 CCTAATGATGATCTTCTATCTACAGGGTAGTTCAGCAATAGAGATTTCTTGATTTA 888  
 QY 326 SerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLys 345  
 DB 889 ACTCCACACCTTCATTTATTTCTCCAGGAAGCACITTCCTGATATATTGGAGAAACACCT 948  
 QY 346 AlaAspPhePheLysArgIleIleGlyLeuLysGluSerSerGluIleCysTyrArg 365  
 DB 949 AAAGAGTTCTTCGAGAAGAAGATCAAGCCATCAGACGCAACGCTTCATGCTGAG 1008  
 QY 366 GluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 385  
 DB 1009 AGGCTCAAGGATATTCCTTGTCTCTTTGTCCAGAAACCCGCAATCTTTCTTATTTA 1068  
 QY 386 MetValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspIleAspPheCysCys 405  
 DB 1069 TGGTTGAAGCTTGACACATCAATGTTGATAATATCAAAATGATTTTGTTCGACG 1128  
 QY 406 LysLeuAlaLysGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsn 425  
 DB 1129 AAGCTAGTTAGTGAGGAGTCTTATCTTATACAGGAGTGGCTCTAGGGGACAGAAAT 1188  
 QY 426 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVal 445  
 DB 1189 TGGTGAGGATATCGATAGGAACCGGCAATCAGTGTCAAGAAATATTTGCAGACTA 1248  
 QY 446 LysSerPheCysGlnArgAsn 452  
 DB 1249 AAAGTTTCTATGATCGTCAT 1269

RESULT 6  
 ARC45930  
 ID AAC45930 standard; DNA; 1533 BP.  
 XX AC  
 XX AAC45930;  
 DT 18-OCT-2000 (first entry)





Db 1313 AAAGTTTCTATGATCGTCAT 1333  
AAC35232  
RESULT 7  
ID AAC35232 standard; DNA; 1535 BP.  
XX AAC35232;  
AC AAC35232;  
DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 9451.  
DE Arabidopsis thaliana  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
PN  
XX  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125785.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130445.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135253.  
PR 24-MAY-1999; 99US-0135623.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 08-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.

```

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 5,59e-82
Score: 1017.50
Percent Similarity: 66.28%
Conservative: 84
Best Local Similarity: 46.60%
Query Match: 42.57%
DB: 21

US-10-019-783-1 (1-461) x AAC35232 (1-1535)

Qy 26 AlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValAluValGluTyrAsnPheAla 45
Db 68 GCGAGCAACGAGGTACCACTGTATACCAACCCCAATGTT-----TGGCGGTTCC--- 118

Qy 46 ArgGlyLysAspGlyLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArg 65
119 -----AAAGGAAACGGTCACAGTAGTCGCGCGCGGTGAGTTCAGTTCGTTGCT 172
Qy 66 TyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHis 85
173 TTTGGGATGTTTAAAACTGCACATGAACAGTGGAAAGACCATTTTGTCCCACTCCC 232
Qy 86 GlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGluAlaValAla 105
233 GCGAGCCCTCCGCCCATTCACACTTCAGAGTTCGCCGGAAGCCGAGAGCGGTGCC 292
Qy 106 AlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAla 125
293 GACGCTGCACGCTCCGCCATCTTACTTACGACCCAGCCCTCGAGTTTCAAGGCT 352
Qy 126 ArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAsp 145
353 AGAAGGCGGTGGCTGAATATTAAACGGAGAACTTCCGACGAAGCTGAAGCGCAGGAT 412
Qy 146 ValPheLeuThrAlaGlyThrGlnAlaIleGluValIleIleProValLeuAlaGln 165
413 GTGTATATCACCGAGGATGTAAACCAAGCCATAGAGATCGTAGATTCTCTTGGCGGA 472
Qy 166 ThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAla 185
473 AATCCATCCGCAACATTTCTTCCAGGCGGGATATCTTCACTACGATGCTCGTGTCT 532
Qy 186 AlaPheAsnLysLeuValAlaArgHisPheAspLeuIleProAspLysGlyTrpGluIle 205
533 GTCTATAGCGGCTCGAGATTCGAAATACGATCTCTCCCGAGAGTGAATGGGAAATC 592
Qy 206 AspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsn 225
593 AATCTCGATGGCTCGAGCGGCTCGGGATGAGATACCGTCGCAATGGTATCATCAAC 652
Qy 226 ProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluVal 245
653 CCCAACAATCCATGTGAAACGCTACACCTACGACCATCTCAACAAGTCCGCGAGATG 712
Qy 246 AlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGly 265
713 GCTAGAAACTCGGTATATGATATATCCGACGAGTATATGATCATGTTGTATATGA 772
Qy 266 SerAlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGly 285
773 GACAGCCCTTTATTCCTCCATGGGAAGTTTGCATCAATAGCTCCGGTGATCAGCTCGGA 832
Qy 286 SerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAsp 305
833 TCCATATCCAAAGGATGGGTCAACCCAGGCTGGAGATTGGCTGGATGCCATGAACGAT 892
Qy 306 ProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 325
893 CCTAATGGTATCTTTGTATCTACAGGGTAGTTTCAAGCAATAGAGATTTCTTGATTTA 952
Qy 326 SerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLys 345
953 ACTCCACAGCCTTCATTTATTCGAGGAAGCACTTCTGTATATATGGAGAAAACACCT 1012
Qy 346 AlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArg 365
1013 AAAGATTTCTTGAGAGAGAGATCAAGCCATGACGACCGCACTCGAGCTTTCATGTGAG 1072
Qy 366 GluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 385
1073 AGGCTCAAGGATATCTTGTCTCTTTGTCCCAAGAAACCCGAATCTTGTCTTATTTA 1132
Qy 386 MetValValLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCys 405
1133 TGGTTGACCTTGACATCAATGTTGAATAATAATCAAAATGATTTTGTATTTTCGACG 1192
Qy 406 LysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsn 425

```

```
Db 1193 AAGCTAGTGTAGTGAGGAGGCTTATCTCTATACGAGGTGGCTTAGGGGAGAGAT 1252
Qy 426 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluAsgVal 445
Db 1253 TGGGTGAGGATATCGATAGGACCGACCAATCAGTGGTACAGAAATATTGACAGACTA 1312
Qy 446 LysSerPheCysGlnArgAsn 452
Db 1313 AAAGTTTCTATGATCGTCAT 1333

RESULT 8
AAC45050
ID AAC45050 standard; DNA; 1481 BP.
AC AAC45050;
XX
XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 45110.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX PD
XX PD
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146385.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
```



PR	09-AUG-1999;	99US-0147935.	Query Match:	36.49%	Indels:	26
PR	10-AUG-1999;	99US-0148171.	DB:	21	Caps:	7
PR	11-AUG-1999;	99US-0148319.	US-10-019-783-1 (1-461) x AAC45050 (1-1481)			
PR	12-AUG-1999;	99US-0148341.	QY	30	LysSerAsnGlyHis-----	AlaAlaAlaAlaValGluTrpAsn
PR	13-AUG-1999;	99US-0148565.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	13-AUG-1999;	99US-0148684.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	16-AUG-1999;	99US-0149368.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	17-AUG-1999;	99US-0149175.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	18-AUG-1999;	99US-0149426.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	20-AUG-1999;	99US-0149722.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	20-AUG-1999;	99US-0149723.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	20-AUG-1999;	99US-0149929.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	23-AUG-1999;	99US-0149902.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	23-AUG-1999;	99US-0149930.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	25-AUG-1999;	99US-0150566.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	26-AUG-1999;	99US-0150884.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	27-AUG-1999;	99US-0151065.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	27-AUG-1999;	99US-0151066.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	27-AUG-1999;	99US-0151080.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	30-AUG-1999;	99US-0151303.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	31-AUG-1999;	99US-0151438.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	01-SEP-1999;	99US-0151930.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	07-SEP-1999;	99US-0152363.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	10-SEP-1999;	99US-0153070.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	13-SEP-1999;	99US-0153758.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	15-SEP-1999;	99US-0154018.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	16-SEP-1999;	99US-0154039.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	20-SEP-1999;	99US-0154779.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	22-SEP-1999;	99US-0155133.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	23-SEP-1999;	99US-0155486.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	24-SEP-1999;	99US-0155659.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	28-SEP-1999;	99US-0156458.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	29-SEP-1999;	99US-0156596.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	04-OCT-1999;	99US-0157117.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	05-OCT-1999;	99US-0157753.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	06-OCT-1999;	99US-0157865.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	07-OCT-1999;	99US-0158029.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	08-OCT-1999;	99US-0158232.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	12-OCT-1999;	99US-0158369.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	13-OCT-1999;	99US-0159293.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	13-OCT-1999;	99US-0159294.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	13-OCT-1999;	99US-0159295.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	14-OCT-1999;	99US-0159329.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	14-OCT-1999;	99US-0159330.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	14-OCT-1999;	99US-0159331.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	14-OCT-1999;	99US-0159637.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	14-OCT-1999;	99US-0159638.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	18-OCT-1999;	99US-0159584.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	21-OCT-1999;	99US-0160741.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	21-OCT-1999;	99US-0160767.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	21-OCT-1999;	99US-0160768.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	21-OCT-1999;	99US-0160770.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	21-OCT-1999;	99US-0160814.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	21-OCT-1999;	99US-0160815.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	22-OCT-1999;	99US-0160980.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	22-OCT-1999;	99US-0160981.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	22-OCT-1999;	99US-0160989.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	25-OCT-1999;	99US-0161404.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	25-OCT-1999;	99US-0161405.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	25-OCT-1999;	99US-0161406.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	26-OCT-1999;	99US-0161359.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	26-OCT-1999;	99US-0161359.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	26-OCT-1999;	99US-0161360.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	26-OCT-1999;	99US-0161381.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	28-OCT-1999;	99US-0161920.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	28-OCT-1999;	99US-0161921.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	28-OCT-1999;	99US-0161933.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
PR	29-OCT-1999;	99US-0162142.	PR	30	-----	AlaAlaAlaAlaValGluTrpAsn
Alignment Scores:						
Pred. No.:			6.85e-69			
Score:			872.00			
Percent Similarity:			60.72%			
Best Local Similarity:			40.63%			
Matches:			1481			
Conservative:			180			
Mismatch:			89			
Mismatches:			148			

Qy 362 IleCysTyrArgGluIleLeuGluAsnLysTyrIleThrCysProHisLysProGluGly 381  
 Db 990 TTTGGTTATTCTAGCTCAAGTACATTCTTAGCTCAGCTTGTACATGAACCCGAGCC 1049  
 Qy 382 SerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluLeuHisAspIle 401  
 Db 1050 TGCACCTTCTATGACCGAGCTTGATTATTCGAGCTTTGTGACATCGAAGACGATCAA 1109  
 Qy 402 AspPheCysCysLysLeuAlaLysGluLysValIleLeuCysProGlySerValLeu 421  
 Db 1110 GACTTTTGCATAAGCTTGCTTAAGAAGAAACCTGCTGTTTACCGAGGATTCATTC 1169  
 Qy 422 GlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerLeuGlnAspGly 441  
 Db 1170 AGTCAGAAGAACTGGTTGAGGCATTCTATCGATATGGAGACTCCGGTATTGGAGGATGCA 1229  
 Qy 442 LeuGluArgValLysSerPheCysGluArgAsnLysLysLysAsnSer-----457  
 Db 1230 TTGGAAAGATTGAAGAGCTTTCGCGATCGCCATTCCAAACAAAGCTCCCTCAAGAC 1289  
 Qy 458 IleAsnGly 460  
 Db 1290 GTCAATGGT 1298  
 RESULT 9  
 AAC39233  
 ID AAC39233 standard; DNA; 1484 BP.  
 XX  
 AC AAC39233;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23860.  
 XX  
 KW Hybridisation assay; Genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EPI033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135363.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139453.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.



```
Qy 312 LysThrLysIleSerThrSerIleThrAsnTyrIleuAsnValSerThrAspProAlaThr 331
Db 841 AACACCAAGGCTTACAAAGTCTCAAGATTTTCCAGATAAACAATAACCCCTCCGACA 900
Qy 332 PheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArg 351
Db 901 GTTATCCAGCGCGTATCTCGACATCTTGAGAAAACCTCCTCAAGAGTTTCTTCGATAAG 960
Qy 352 IleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLys 371
Db 961 AGGCAGAGTTTTCGAAAGATAAAGTAGAATTTGGTTATTCTAAGTCGAGTACATTCCT 1020
Qy 372 TyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeu 391
Db 1021 AGCCTCACTTCTCATCAAGAACCCGAACTGACCTTCTATGACCGAGCTTGATTA 1080
Qy 392 HisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLysGluGlu 411
Db 1081 TCGAGCTTTGTGACATCGAAGACGATCAGACTTTTGCATTAAGCTTGCTAAGAAGAA 1140
Qy 412 SerValIleLeuCysProGlySerValLeuGlyMetGluAsnThrValArgIleThrPhe 431
Db 1141 AACCTCGTCTGTTTACCAGGAGTTGCATTCAGTCAGAAGAACTGGTTGAGCGCATTCATC 1200
Qy 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArg 451
Db 1201 GATATGGAGACTCCGGTATTCGAGATGTCATTGGAAGATTAAGAGACTTCTCGGATCGC 1260
Qy 452 AsnLysLysLysAsnSer-----IleAsnGly 460
Db 1261 CATTCCAAACAAAAGCTCCCTCCCTCAAGACGCTCAATGGT 1299
RESULT 10
AAC45003
ID AAC45003 standard; DNA; 1584 BP.
XX AC AAC45003;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 44934.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway;
XX KW metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 09-MAR-1999; 99US-0123180.
XX PR 23-MAR-1999; 99US-0123548.
XX PR 25-MAR-1999; 99US-0125788.
XX PR 29-MAR-1999; 99US-0126264.
XX PR 01-APR-1999; 99US-0126785.
XX PR 06-APR-1999; 99US-0127462.
XX PR 08-APR-1999; 99US-0128234.
XX PR 16-APR-1999; 99US-0128714.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 04-MAY-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144634.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
```



Db 699 ATGGTCGTTCCGATGAAGTATATAGATGGTCCGTATTCGGAGTAATCATTTGTTCCA 758  
Qy 272 MetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrp 291  
Db 759 ATGGGAAATTCCTCGATAGTACCACTGATTAACCTCGGTCATATCGAAGAGTGG 818  
Qy 292 IleValProGlyTrpArgLeuGlyTrpValAlaValTyrrAspProThrLysIleLeuGlu 311  
Db 819 ATTCTCCAGATCGCAACGGCTGCTCGGTACACCATTAATGGTGTCTTTAGA 878  
Qy 312 LysThrLysIleSerThr-----SerIleThrAsnTyrLeuAsnValSerThrAspPro 393  
Db 879 TCCCAAGAGTCCCGGCTTTAAAGCTCGGAAGAATTTCTGGAGATTAATCTCCAAACCA 938  
Qy 330 AlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePhe 349  
Db 939 CCAACTGTTCATCCAGCGGCTATTCCCAACATCTTGGAGAAATCTCTCAAGATTTTTC 998  
Qy 350 LysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGlu 369  
Db 999 GAGAGAGCGGGATTTTCTAAAGACAAAGTGGATTTGGATATTTAAAGCTCAAGAAC 1058  
Qy 370 AsnLysTyrIleThrCysProHisLysProGlyGlySerMetPheValMet---ValLys 388  
Db 1059 ATACCAACTCTCACCTGCTACATGAACCTGAATCATGCACCTTCTT-ATGGAGACCAAG 1117  
Qy 389 LeuAsn---LeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeu 407  
Db 1118 CTGAGCCCATTCATTTTGTGA---GACATTAAGACCATCAGACTTCTCGAGGAAGCTT 1174  
Qy 408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427  
Db 1175 GCTAAGGAAGAAATCTCGTTGTTTACAGGAATTCATTGCTGCAAGATAACTGGTTG 1234  
Qy 428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447  
Db 1235 AGGCATTCATCGACATGGAGACTCCCAAGATGGAGGATGCAATTTGAAGATTAAAGGC 1294  
Qy 448 PheCysGlnAlaGAsn-----LysLysLysAsnSerIle 458  
Db 1295 TTCGTGTAAGCCATTCGTTATAGTCAAGCTTCTTCACTCAAGAGCTCAATGGGCTC 1354  
Qy 459 Asn 459  
Db 1355 AAC 1357  
RESULT 11  
AAC46309  
ID AAC46309 standard; DNA; 1101 BP.  
XX AC AAC46309;  
XX AC  
DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 49668.  
XX XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX OS Arabidopsis thaliana.  
XX XX  
PN EP1033405-A2.  
XX XX  
PD 06-SEP-2000.  
XX XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999;  
PR 01-APR-1999;  
PR 06-APR-1999;  
PR 08-APR-1999;  
PR 16-APR-1999;  
PR 19-APR-1999;  
PR 21-APR-1999;  
PR 23-APR-1999;  
PR 28-APR-1999;  
PR 30-APR-1999;  
PR 04-MAY-1999;  
PR 05-MAY-1999;  
PR 06-MAY-1999;  
PR 07-MAY-1999;  
PR 11-MAY-1999;  
PR 14-MAY-1999;  
PR 14-MAY-1999;  
PR 14-MAY-1999;  
PR 18-MAY-1999;  
PR 19-MAY-1999;  
PR 20-MAY-1999;  
PR 21-MAY-1999;  
PR 24-MAY-1999;  
PR 25-MAY-1999;  
PR 27-MAY-1999;  
PR 28-MAY-1999;  
PR 01-JUN-1999;  
PR 03-JUN-1999;  
PR 04-JUN-1999;  
PR 07-JUN-1999;  
PR 08-JUN-1999;  
PR 10-JUN-1999;  
PR 10-JUN-1999;  
PR 14-JUN-1999;  
PR 16-JUN-1999;  
PR 16-JUN-1999;  
PR 17-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 18-JUN-1999;  
PR 21-JUN-1999;  
PR 22-JUN-1999;  
PR 23-JUN-1999;  
PR 24-JUN-1999;  
PR 28-JUN-1999;  
PR 29-JUN-1999;  
PR 30-JUN-1999;  
PR 01-JUL-1999;  
PR 01-JUL-1999;  
PR 02-JUL-1999;  
PR 06-JUL-1999;  
PR 09-JUL-1999;  
PR 12-JUL-1999;  
PR 13-JUL-1999;  
PR 14-JUL-1999;  
PR 15-JUL-1999;  
PR 16-JUL-1999;  
PR 16-JUL-1999;  
99US-0126785;  
99US-0127462;  
99US-0128234;  
99US-0128714;  
99US-0129845;  
99US-0130077;  
99US-0130449;  
99US-0130510;  
99US-0130891;  
99US-0131449;  
99US-0132048;  
99US-0132407;  
99US-0132484;  
99US-0132485;  
99US-0132486;  
99US-0132487;  
99US-0132863;  
99US-0134256;  
99US-0134218;  
99US-0134219;  
99US-0134221;  
99US-0134370;  
99US-0134768;  
99US-0134941;  
99US-0135124;  
99US-0135353;  
99US-0135629;  
99US-0136021;  
99US-0136392;  
99US-0136782;  
99US-0137222;  
99US-0137528;  
99US-0137502;  
99US-0137724;  
99US-0138094;  
99US-0138540;  
99US-0138847;  
99US-0139119;  
99US-0139452;  
99US-0139453;  
99US-0139492;  
99US-0139454;  
99US-0139455;  
99US-0139456;  
99US-0139457;  
99US-0139458;  
99US-0139459;  
99US-0139460;  
99US-0139461;  
99US-0139462;  
99US-0139463;  
99US-0139750;  
99US-0139763;  
99US-0139817;  
99US-0139899;  
99US-0140353;  
99US-0140354;  
99US-0140695;  
99US-0140823;  
99US-0140991;  
99US-0141287;  
99US-0141842;  
99US-0142154;  
99US-0142055;  
99US-0142390;  
99US-0142920;  
99US-0142977;  
99US-0143542;  
99US-0143624;  
99US-0144005;  
99US-0144085;  
99US-0144086;



Db	559	ATATCAAGGGATGGAAGTCCAGGATGGCGAATCGTGTGGCTCACGCTACATGATCTA	618	99US-0132407
Qy	307	ThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer	326	99US-0132484
Db	619	GACGGTGCTTCAGAAACACCAAGGCTTCACAGCTGCTCAAGATTTCCTCCAGATAAAC	678	99US-0132486
Qy	327	ThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAla	346	99US-0132487
Db	679	AATAACCCCTCCGACAGTATCCAGCGGCTATTCCTGCATCTTTGGAGAAACCTCTCAA	738	99US-0132488
Qy	347	AspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGlu	366	99US-0132489
Db	739	GAGTCTTTTGAATGAAGGAGGAGAGTTCCTGAAAGATAAAGTAGAATTTCGTTATCTTAAG	798	99US-0132490
Qy	367	IleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet	386	99US-0132491
Db	799	CTCAAGTACATTCCTAGCCTCACTTGTACATGAACCCGGAAGCCTCACCTTCTTATGG	858	99US-0132492
Qy	387	ValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLys	406	99US-0132493
Db	859	ACCGAGCTTGATTTATCGAGCTTTTCGACATCGAAGACGATCAAGACTTTTGGCAATAAG	918	99US-0132494
Qy	407	LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTyr	426	99US-0132495
Db	919	CTTGCTAAGAGAAACCTCTGCTTTTACCAGGAGTGCATTTCAGTCAGAGAACTGG	978	99US-0132496
Qy	427	ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLys	446	99US-0132497
Db	979	TTGAGGCAATCTATCGATATGAGACTCCGGTATTCGAGATGCATTTCGAAAGATTGAAG	1038	99US-0132498
Qy	447	SerPheCysGluArgAsnLysLysAsnSer-----IleAsnGly	460	99US-0132499
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132500
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132501
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132502
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132503
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132504
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132505
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132506
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132507
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132508
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132509
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132510
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132511
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132512
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132513
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132514
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132515
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132516
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132517
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132518
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132519
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132520
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132521
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132522
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132523
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132524
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132525
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132526
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132527
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132528
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132529
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132530
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132531
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132532
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132533
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99US-0132534
Db	1039	AGCTTCTGCATCGCCATTCACAAACAAAGCTCCCTCCAAAGAGCTCAATGGT	1092	99



```
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146398.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 7,35e-62 Length: 1182
Score: 792.50 Matches: 150
Percent Similarity: 66.57% Conservative: 75
Best Local Similarity: 44.38% Mismatches: 108
Query Match: 33.16% Indels: 5
DB: 21 Gaps: 2

US-10-019-783-1 (1-461) x AAC45560 (1-1182)
Qy 127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspVal 146
Db 11 AGTGGCGTAGCAGAGTATCTAAACCAAGGCTCTCCAAAGAGCTTACCAGAGTACGCTG 70
Qy 147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThr 166
Db 71 TTTATGACTCTGGGATGCAACCAAGCTATTGAGCTCGCGTAGGACCTCGCTCAACCG 130
Qy 167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaA 186
Db 131 AAA--GCCAACGTTTGTCTCCGAGTCCCGGCTTCCCATGGGACCTAGTCCGCTCATC 187
Qy 187 PheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp 206
Db 188 TACAAGAACCTTGAGTCCGCCACTATAATTCCTCCAGAAAGAACTTGAATCGAC 247
Qy 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrAlaMetValIleIleAsnPro 226
Db 248 TTGTATAGGCTCCGAGCGGCTCGTGGACGAGACACATTTGCCATATTTATATCAACCCC 307
Qy 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246
Db 308 CACACCCCAATGGTAACACTTACTCCGAGGCTCATCTCAACAGCTGGCTGAACCTGGCT 367
Qy 247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 266
Db 368 AAGGAACCTAAGATTATGTGGTTTCTGACGAGGTTTATAGATGGACACTTTTGTAGT 427
Qy 267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer 286
Db 428 AACCCTTTGTCTCTATGGGAAATTTCTCGATCGTACCATCGAGTGGTGTACACTCGGATCC 487
Qy 287 LeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspPro 306
Db 488 ATATCAAGGGATGGAAGATCCAGGATGCCAGACTGGTGGCTACGCTACGATGATCTA 547
Qy 307 ThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 326
Db 548 GACGGTGTCTTCAGAAACACCAAGGCTCTTACAGCTGCTCAAGATTTTCTCCAGATAAC 607
Qy 327 ThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAla 346
Db 608 AATAACCCCTCCGACAGTATCCAGGCGGCTATTCTGACATCTTGGAGAAACTCTCTCAA 667
```

```
OY 347 AspPhePheLysArgIleileGlyLeuLeuLysGluSerSerGluIleCysTyrArgGlu 366
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 668 GAGTTTTTGTATAGAGCAGCAGATTCTGAAAGATAAAGTAGAATTGGTTATCTTAAG 727
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
OY 367 IleLysGluAenLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet 386
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 728 CTCAGTAGTACATTCTAGCCTCACITGCTACATGAACCCGAAGCCTGCACCTCTCTTAGG 787
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
OY 387 ValLysLeuAenLeuHisLeuLeuGluCluIleHisAspAspIleAspPheCysCysLys 406
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 788 ACCGAGCTTGATTATCGAGCTTGTGGACATCGAAGACATCAAGACTTTTGCATTAAG 847
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
OY 407 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrp 426
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 848 CTTGCTAAAGAAGAAACCTCGTCTTTTACAGGATTGCATTCAGTCAGACGAACATCGG 907
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
OY 427 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLys 446
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 908 TTGAGGCATTCTATCGATATGAGACTCCCGTATTGGAGGATGCATTGGAAAGATTGAAG 967
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
OY 447 SerPheCysGlnArgAsnLysLysAsnSer-----IleAsnGly 460
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 968 ACCTTCGCGATCGCATCCACGAAAGAACTCCCTCAAGACGTCATGGT 1021
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

RESULT 13
AAC36514
ID AAC36514 standard; DNA; 1184 BP.
XX
AC AAC36514;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14092.
KW Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence, ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
```

```
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1.69e-61 Length: 1184
Score: 788.50 Matches: 149
Percent Similarity: 66.57% Conservative: 76
Best Local Similarity: 44.08% Mismatches: 108
Query Match: 32.99% Indels: 5
DB: 21 Gaps: 2

US-10-019-783-1 (1-461) x AAC36514 (1-1184)
QY 127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyr-LysLeuSerAlaAspVal 146
Db 11 AGTGGCGTAGCAGATATCTAAACCAAGGCTTCCAAAGAGCTTACAGCAGATGACGTG 70
QY 147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThr 166
Db 71 TTTATGACCCCTGGATGCAACCAAGCTATTGAGCTCGCGGTAGACATTCTCGCTAACCG 130
QY 167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 186
Db 131 AAA---GCCAACGCTTTTGCTTCGAGTCCCGGCTCCCATGGGACCTAGTCCGCTCCATC 187
QY 187 PheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAsp 206
Db 188 TACAAGAACCTTGAGTCCGCCACTATATTCCTTCCGAAAGAACTTTGAAATCGAC 247
QY 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 226
Db 248 TTTGATAGCGTCCGAGCGCTCGTGGACGAGACACATTTGCCATATTTATATCAACCCC 307
QY 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246
Db 308 CACACCCCNATGTAACACCTACTCCGAGGCTCATCTCAACAGCTCGCTGAATGGCT 367
QY 247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 266
Db 368 AAGGAACTCAAGATTATGTGGTTTCTGACAGGTTTGTAGATGGACACTTTGGGAGT 427
QY 267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer 286
Db 428 AACCCCTTTTGCCCTATGGGAAATTTCTCGTACGATCGTACCAAGTGTGTACACTCGGATCC 487
QY 287 LeuSerLysSerTyrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspPro 306
Db 488 ATATCAAGGGATGGAAGATCCAGGATGCGNACTGGTGGCTACCGCTACATGATCTA 547
QY 307 ThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 326
Db 548 GACGGTGCTTTCAGAAACACCAAGGCTTACAAGCTGCTCAGGATTTTCTCCAGATAAAC 607
QY 327 ThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAla 346
Db 608 AATACCCCTCCGACAGTTATTCAGGGCGGTATCTCTGACATCTTGGAGAAACTCTCTAA 667
QY 347 AspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGlu 366
Db 668 GAGTTCTTCGATAAGAGCGCAGAGATTTCCTGAAAGATAAAGTAGAATTTGGTTATTCTAAG 727
QY 367 IleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet 386
Db 728 CTCGAGTACATTCTCTAGCCTCACTTCTGCTACATCAACCCGAGCCCTGACCTCTTATCG 787
QY 387 ValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLys 406
Db 788 ACCGAGCTTGATTATCGAGCTTTGGGACATCGAGACATCGAGACATTTTGCATATAG 847
```

QY 407 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrp 426  
DB 848 CTTCCTAAGAGAGAAACCTCGTCTTTACAGGGATTGCTTCAGTCAGAGAACTGG 907  
QY 427 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLys 446  
DB 908 TTGAGGCATCTCATGATATGGAGACTCCGGTATTGGAGGATGATTCGGAAGATTGAAG 967  
QY 447 SerPheCysGlnArgAsnLysLysAsnSer-----IleAsnGly 460  
DB 968 AGCTTCTCGATGCGCATTCACAAAGAAAGCTCCCTCAAGACGTCATGGT 1021  
RESULT 14  
ID ABN98840/c  
ID ABN98840 standard; DNA; 858 BP.  
AC ABN98840;  
XX  
XX  
XX  
XX 01-AUG-2002 (first entry)  
XX Arabidopsis thaliana expressed polynucleotide SEQ ID NO 608.  
XX Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;  
KW disease; crop; thale cress; tolerance factor; insect; pathogen;  
KW nutrition; ds.  
XX  
XX Arabidopsis thaliana.  
XX  
XX US2002023281-A1.  
XX  
XX  
XX 21-FEB-2002.  
XX  
XX 26-JAN-2001; 2001US-0770445.  
XX  
XX 27-JAN-2000; 2000US-178472P.  
XX  
XX (GORL/) GORLACH J.  
XX (ANY/) AN Y.  
XX (HAMI/) HAMILTON C M.  
XX (PRIC/) PRICE J L.  
XX (RAIN/) RAINES T M.  
XX (YUY/) YU Y.  
XX (RAME/) RAMEKA J G.  
XX (PAGE/) PAGE A.  
XX (MATH/) MATHW A V.  
XX (LEDF/) LEDFORD B L.  
XX (WOES/) WOESSNER J P.  
XX (HAAS/) HAAS W D.  
XX (GARC/) GARCIA C A.  
XX (KRIC/) KRICKER M.  
XX (SLAT/) SLATER T.  
XX (DAVI/) DAVIS K R.  
XX (ALLE/) ALLEN K.  
XX (HOFF/) HOFFMAN N.  
XX (HURB/) HURBAN P.  
XX  
XX Goriach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;  
PI Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Krieker M, Slater T, Davis KR, Allen K, Hoffman N,  
PI Hurban P;  
XX  
XX WPI; 2002-400781/43.  
XX  
XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,  
PT producing compositions that modulate the expression or function of its  
PT encoded protein, and mapping functional regions of protein  
XX  
XX Claim 1; SEQ ID NO 608; 49pp + Sequence Listing; English.  
XX  
XX The invention relates to an Arabidopsis thaliana nucleic acid (I)  
CC comprising a sequence capable of hybridising under stringent conditions  
CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),

CC given in the specification or its fragment. A polypeptide (II) encoded by  
CC (II), a transgenic plant (III) comprising an exogenous nucleic acid or a  
CC genetically modified cell (IV) comprising an exogenous nucleic acid, is  
CC useful for screening a candidate agent for its biological effect. (I) is  
CC useful in identifying homologous or related genes, in producing  
CC compositions that modulate the expression or function of its encoded  
CC protein, mapping functional regions of the protein and in studying  
CC associated physiological pathways. (I) is also useful for the genetic  
CC manipulation of cells, particularly plant cells. (I) is also useful in  
CC screening assays of various plant strains to determine the strains that  
CC are best capable of withstanding a particular disease or environmental  
CC stress. (II) and (III) are useful for screening of biologically active  
CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical  
CC pathways. The screened agents are useful in improved methods of treating  
CC crops to prevent or treat disease. (II) are also useful in screening  
CC programs to identify agents that mimic or enhance the action of tolerance  
CC factors. Such agents are useful in improved methods of treating crops to  
CC enhance their tolerance to environmental stress. (I) is also useful  
CC for enhancing or inhibiting production of a biosynthetic product in a  
CC plant. (III) is useful for identifying other mediators that may induce  
CC expression of proteins of interest, for establishing the extent to which  
CC any specific insect and/or pathogen is responsible for damage to a  
CC particular plant, for identifying other mediators that enhance or induce  
CC tolerance to environmental stress, for identifying factors involved in  
CC biosynthetic pathways of nutritional, commercial, or medicinal value and  
CC for identifying productions of nutritional, commercial or medicinal  
CC value. (IV) is useful in the study of genetic function and regulation,  
CC for alteration of the cellular metabolism and for screening compounds  
CC that may affect the biological function of the gene or gene products.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=999909770445.  
XX  
XX SQ Sequence 858 BP; 254 A; 188 C; 165 G; 251 T; 0 other;

## Alignment Scores:

Pred. No.: 4,06e-49 Length: 858  
Score: 649.00 Matches: 118  
Percent Similarity: 70.6% Conservative: 41  
Best Local Similarity: 52.44% Mismatches: 66  
Query Match: 27.15% Indels: 0  
DB: 24 Gaps: 0

US-10-019-783-1 (1-461) x ABN98840 (1-858)

QY 230 CysGlySerValTySerTyAspHisLeuAlaGluValAlaArgLysLeu 249  
DB 858 TGTGGATGTTTCTCTCGCCACATCTTCAAAAGATTGCAGACGCGCTTGAAGCTT 799  
QY 250 GlyIleLeuValIleAlaAspGluValTyGlyLysLeuValLeuGlySerAlaProPhe 269  
DB 798 GGTATACITGTGATCGCAGAGAGTCTATGACCATTTTGGCTTTGGGATAAACCAATT 739  
QY 270 IleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLys 289  
DB 738 GTGTCCATGGCAGAGTTTGCAGAGTAGTGCCTGTATTGTCTTAGTGCCTATATCTAAA 679  
QY 290 SerTPIleValProGlyTIPArgLeuGlyTIPValAlaValTyAspProThrLysIle 309  
DB 678 AGATGGTTTGTTCCTGGATGGAGACTTGTGGATGGTGACTTGTGACCCATCGGCATC 619  
QY 310 LeuGluLysThrLysIleSerThrSerIleThrAsnTyLeuAsnValSerThrAspPro 329  
DB 618 ATGAAGATCTCGGGTTTGTTCAGACTCTTATCAATGTGTCAACATGTCACGATCCT 559  
QY 330 AlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePhe 349  
DB 558 GCAACGTTTATTCAGGAGCATGCTGATATCATTCATGGATACAAAGAGAGTCTTC 499  
QY 350 LysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyArgGluIleLysGlu 369  
DB 498 TCATCAAAACCTTGAATGTGAAAAAATGTCAGAGATTGTTTATGAGAGCTTATGAAG 439





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 19:50:55 ; Search time 2809.28 Seconds  
(without alignments)  
3988.346 Million cell updates/sec

Title: US-10-019-783-1  
Perfect score: 2390  
Sequence: 1 MVHOSNGHGEAAAAANGS.....LERVKSFCQKKNKNSINGC 461

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10019783/runat\_07112003\_161643\_19848/app.query.fasta\_1.1358  
-DB=EST -OEW=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LQOPEC=0 -LQOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THRESHOLD=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZES=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10019783 -CGN\_1\_1\_4125@runat\_07112003\_161643\_19848 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1376.5	57.6	1143	11	AY104359 Zea mays
2	1177	49.2	770	10	BF268125 HV CBA001
3	1166.5	48.8	813	14	CB525593 OSJNEa50
4	1070	44.8	657	14	CA015831 HT1405f
5	1024.5	42.9	891	14	CD437584 ELOIN0502
6	1020.5	42.7	837	14	CB668395 OSJNEd16D
7	1020	42.7	843	14	CB668396 OSJNEd16D
8	1020	42.7	878	14	CB643965 OSJNEb04P
9	1020	42.7	902	14	CB618033 OSJNEa01J
10	1019	42.6	849	14	CB671005 OSJNEa04F
11	1015	42.5	832	14	CB652840 OSJNEc03E
12	1001	41.9	883	14	CB646007 OSJNEb08C
13	992.5	41.5	809	14	CB618032 OSJNEa01J
14	981	41.0	619	10	BG313019 WHE0091.D
15	975	40.8	606	13	BQ466157 HT01121T
16	960.5	40.2	792	14	CB670913 OSJNEa04D
17	938	39.2	836	14	CB670014 OSJNEe02O
18	912	38.2	565	10	BF065437 HV_CEB001
19	906	37.9	595	14	CA599545 waw1c.pk0
20	897	37.5	816	14	CB668541 OSJNEd16H
21	887.5	37.1	745	14	CB678552 OSJNEe16L
22	886.5	37.1	751	14	CB668540 OSJNEd16H
23	874.5	36.6	740	14	CB646006 OSJNEb08C
24	871	36.4	743	14	CB675132 OSJNEe10K
25	857.5	35.9	604	9	AW257959 68704G08
26	848	35.3	683	14	CD210710 HSL_49_A1
27	845	35.4	668	14	CD230874 SSI_20_D1
28	830.5	34.7	811	14	CB671004 OSJNEa04F
29	825	34.5	744	14	CB670914 OSJNEa04D
30	823.5	34.5	806	14	CB643964 OSJNEb04P
31	815	34.1	734	14	CA401705 ELOIN0424
32	814	34.1	664	14	CB924613 ABA1_1.F0
33	811.5	34.0	577	10	BE230084 99A3319.R
34	811	33.9	806	14	CD439867 ELOIN0530
35	807	33.8	763	13	BQ967714 QHB39K17
36	807	33.8	791	13	BQ969834 QHB39K17
37	806	33.7	659	14	CD231700 SSI_23_A0
38	806	33.7	670	14	CD236444 SSI_34_B1
39	802.5	33.6	562	13	B0050754 1111033B0
40	801	33.5	650	10	BE598422 P11_82.G0
41	798	33.4	660	14	CD210858 HSL_66_A0
42	797	33.3	777	13	BQ915490 QHB14M12
43	796	33.3	624	10	BE598586 P11_84.G0
44	793.5	33.2	596	10	BE365047 P11_23_B1
45	792	33.1	481	10	BF200495 WHE2258_G

ALIGNMENTS

RESULT 1	AY104359	PC0115235	1143 bp	mrna	linear	HTC 16-OCT-2002
LOCUS	AY104359					
DEFINITION	Zea mays					
ACCESSION	AY104359					
VERSION	AY104359.1	GI:21207437				
KEYWORDS	HTC.					
SOURCE	Zea mays					
ORGANISM	Zea mays					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1	(bases 1 to 1143)				





source

1..770

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="C116155 (M1a13)"

/db\_xref="taxon:112509"

/clone="HV\_CEA0019P10f"

/tissue\_type="seedling green leaf"

/lab\_host="TUC121"

/clone\_lib="Hordeum vulgare seedling green leaf EST library HVCNDA0004 (Blumeria challenged)"

/notes="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI; C.I. 16155 (M1a13) plants were greenhouse grown in the R Wise lab at Iowa State University, Ames, IA; 7 day old green seedlings were challenged with isolate A27 (AvrWl13) of Blumeria graminis f. sp. hordei, and leaves were harvested 20 and 24 hr post-inoculation and snap frozen; uninoculated leaves were harvested 20 hr post-inoculation (Wei, Wise). In the TJ Clouse lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids (Choi, Clouse). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phased value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Clouse TJ, Wing R, Kleinbols A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 164 a 247 c 216 g 143 t

ORIGIN

Alignment Scores:

Prod. No.: 5,27e-115 Length: 770

Score: 1177.00 Matches: 230

Percent Similarity: 94.49% Conservative: 10

Best Local Similarity: 90.55% Mismatches: 13

Query Match: 49.25% Indels: 1

DB: 10 Gaps: 0

US-10-019-783-1 (1-461) x BF68125 (1-770)

QY 93 AlaPheArThrAlaValGluAlaGluAspAlaValAlaAlaLeuArGThrGlyGln 112

Db 3 GCCTTCGACGCGCGTTCAGGCCGAGGACGCGCGCGCGCGTGCACCGGCAG 62

QY 113 PheAsnCysTyAlaAlaGlyValGlyLeuProAlaAlaArGSerAlaValAlaGluHis 132

Db 63 TTCACCTGCTACCCGCGCGCGTTCGCGTCCCGCGGACGAGCGCGTGCAGAGCAC 122

QY 133 LeuSerGlnGlyValProGlyTyrysLeuSerAlaAspValPheLeuThrAlaGlyGly 152

Db 123 CTGTCGACGGGTCGCGTACAGGTATCGCGCAGCAGCGTCTTCTCCACGCGCGCGG 192

QY 153 ThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeu 172

Db 183 ACCGAGCGCATCGAGGTATATATCCGCTGCTGCGCCACGCGCGCGCGCAATCTG 242

QY 173 LeuProArGProGlyTyProAsnTyTyGluAlaArgAlaAlaPheAsnLysLeuGluVal 192

Db 243 CTCGCCAGCCAGGCTATCCAACTACGAGGCGCGCGCGGTTTCACAGGCTGGAGTC 302

QY 193 ArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSer 212

303 CGGCATTTTCGACCTCATCCCGACAAAGGGTGGGAGATCGACATCGCTCGAATCC 366

QY 213 IleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySer 232

Db 363 ATCGCCGCAACAGACACACACCCCATGGTCTCATPAACCCCAACACCCGTCGGGAGC 422

QY 233 ValTySerTyAspHisLeuAlaLysValAlaGluValAlaAlaArgLysLeuGlyIleLeu 252

Db 423 GTTTACTCTGACGACCATCTGTCCAAGGTTCGCGAGGTGGGAAAGGCTCGGAATATTG 482

QY 253 ValIleAlaAspGluValTyGlyLysLeuValLeuGlySerAlaProPheIleProMet 272

Db 483 GTGATTCCTGACGAGGTATACGCAAGTGTGTTCGGCAGCGCGCGCTTCATCCCAATG 542

QY 273 GlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIle 292

Db 543 CGAGTGTTCGGGCACATCACCCCTGTGTCTCCATAGGTCTCTGTCTAAGTCATGGATA 602

QY 293 ValProGlyTrpArgLeuGlyTrpValAlaValTyAspProThrLys-IleLeuGluGly 312

Db 603 GTGCTTCGATGCGCGTTCGATGGGTACGGGTACGAAACCCCAAGAGGATCTTCGCGGA 662

QY 312 SThrLysIleSerThrSerIleThrAsnTyrysLeuAsnValSerThrAspProAlaThrPh 332

Db 663 AACTAAGAACTTTACATCCATTCAGAACTTAACTTAACTTCTCGACAGACCCACACACTT 722

QY 332 etValGlnGluAlaLeuProLysIleLeuGluAsnThrLys 345

Db 723 TATTCAGCGCAGCTCTTCTCAGAAATTTTGGAAACACCAAG 762

RESULT 3

LOCUS CB625993 813 bp mRNA linear EST 08-APR-2003

DEFINITION OSIIea5015.f OSIIea Oryza sativa (indica cultivar-group) cDNA

ACCESSION Clone OSIIea5015.5, mRNA sequence.

VERSION CB625993

KEYWORDS CB625993.1 GI:29620982

SOURCE EST.

ORGANISM Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 813)

Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe oryzae

Unpublished

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA

Tel: 520 826 3967

Fax: 520 621 9288

Email: <http://genome.arizona.edu>

PCR Primers

FORWARD: gta aaa cga cgg cca gtc

BACKWARD: gga aac agc tat gac cat g

Plate: 15 row: O column: 15

Seq primer: gta aaa cga cgg cca gtc.

Location/Qualifiers

1..813

/organism="Oryza sativa (indica cultivar-group)"

/mol\_type="mRNA"

/cultivar="IR36"

/db\_xref="taxon:39946"

/clone="OSIIea5015"

/tissue\_type="Leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_lib="OSIIea"

FEATURES

source

/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:

XhoI; Lesion Mimic SPL 11"

BASE COUNT 194 a 235 c 191 g 193 t

ORIGIN

# Alignment Scores:

Pred. No.: 7,72e-115 Length: 813  
Score: 1166.50 Matches: 219  
Percent Similarity: 90.41% Conservative: 26  
Best Local Similarity: 80.81% Mismatches: 25  
Query Match: 48.81% Indels: 1  
DB: 14 Gaps: 1

US-10-019-783-1 (1-461) x CB625993 (1-813)

QY 68 l1eSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87  
DB 1 ATCAGCGCCAGGTCCTACCTACCGCGCGCGCGCGCGCTCTCGCTCGCGCGCGGAC 60  
QY 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107  
DB 61 CCTCCGTGTTCCCGAGTTCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120  
QY 108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127  
DB 121 CTCCTCGCGCGGACTTCACTGCTACCGCGCGCGCGCGCGCTCTCGCGCGCGCG 180  
QY 128 AlavalAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147  
DB 181 GCTGTGGCAGATCATTTGTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
QY 148 LeuThrAlaGlyGlyThrGluAlaLeuGluVallelleProValLeuAlaGlnThrAla 167  
DB 241 CTAACCGCTGGAGGAACCTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 297  
QY 168 GlyAlaAsnleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187  
DB 298 GCACAAACATATTCTCTCTAGACCGGTCACCAACTATGAAGCTCGAGCGCGGTC 357  
QY 188 AsnLysLeuGluValAlaArgHisPheAspLeuLeuProAspLysGlyTyrGluAlaAspIle 207  
DB 358 AACAACTTGAAGTTCGTCACCTTGTATCTTATCTGAGAGGGCTGGGAGATTGACCTT 417  
QY 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetVallelleAsnProAsn 227  
DB 418 AACTCCCTAGATCTATTGGCGACAAAGCACTACTGCGATAGTCATCATTAATCCCAAT 477  
QY 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247  
DB 478 AATCCATGCGGGAATGTGTACACTTACGAGCATTTATCCAAAGGTGGCAGAGTAGCAAGG 537  
QY 248 LysLeuGlyIleLeuVallelleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267  
DB 538 AAGCTTGGGATATTGGTATTAATGATGAGGTGATGATGATGATGATGATGATGATG 597  
QY 268 PropheileProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287  
DB 598 CCATTGTGCCAATGGTGTCTTGGGACATCGTACCAATATTAAACCATAGGATCGCTA 657  
QY 288 SerLysSerTyrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThr 307  
DB 658 TCAAGAGGTGATAGTGGCGGATGGGACTTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 717  
QY 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327  
DB 718 AAGACTCTACAAGAAACCAAGATTCAACATTAATTAATTAATTAATTAATTAATTAAT 777  
QY 328 AsproAlaThrPheValGlnGluAlaLeuPro 338  
DB 778 GATCCAGCAACTTTTATTTCAGGAGGCTCTTACCG 810

RESULT 4  
CA015631

LOCUS CA015631 657 bp mRNA linear EST 23-OCT-2002  
DEFINITION HT14013r HT Hordeum vulgare subsp. vulgare cDNA clone HT14015  
5-PRIME, mRNA sequence.  
ACCESSION CA015631  
VERSION CA015631.1 GI:24292975  
KEYWORDS EST  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
; Triticeae; Hordeum.  
REFERENCE 1 (bases 1 to 657)  
AUTHORS Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner, A.  
TITLE Barley ESTs from germinating seeds  
JOURNAL Unpublished  
COMMENT Contact: Stein Nils  
Molecular Markers Group, Department Genbank  
Institute of Plant Genetics and Crop Plant Research (IPK)  
Corrensstr. 3, 06466, Gatersleben, Germany  
Tel: 039482-5522  
Fax: 039482-5595  
Email: stein@ipk-gatersleben.de  
Insert Length: 657 Std Error: 0.00  
Plate: 14 row: 0 column: 15  
Seq primer: M13rev.  
FEATURES  
Location/Qualifiers  
1..657  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="barke"  
/db\_xref="GABI:264169"  
/db\_xref="taxon:112509"  
/clone="HT14015"  
/tissue\_type="endosperm early"  
/dev\_stage="0-16 hours after imbibition"  
/lab\_host="XL10-Gold"  
/clone\_lib="HT"  
/note="Vector: pBluescript SK+; Site\_1: EcoRI (5'-end of cDNA); Site\_2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."  
BASE COUNT 140 a 209 c 187 g 121 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.36e-104 Length: 657  
Score: 1070.00 Matches: 205  
Percent Similarity: 98.14% Conservative: 6  
Best Local Similarity: 95.35% Mismatches: 4  
Query Match: 44.77% Indels: 0  
DB: 14 Gaps: 0  
US-10-019-783-1 (1-461) x CA015631 (1-657)  
QY 113 PheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHis 132  
DB 12 TTCACTGTACCCCGCGCGCTCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 71  
QY 133 LeuSerGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGly 152  
DB 72 CTCTCGCAGGCGTGGGTACATGCTATCGCGCGCGCGCGCTTCTCTACCGCGCGGG 131  
QY 153 ThrGlnAlaIleGluVallelleProValLeuAlaGlnThrAlaGlyAlaAsnleLeu 172  
DB 132 ACCAGCGCATCGAGTGTATATCCCGGTGCTGGCCAGCGCGCGCGCGCGCGCG 191  
QY 173 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluVal 192

192 CCCCCAGCCAGGCTACCCAACTACGAGGCGCGCGGCTTCAACAGGCTGAGGTC 251  
 QY 193 AtgHisPheAspLeuIleProAspLysGlyTyrGluIleAspSerLeuGluSer 212  
 Db 252 CGGCAATTCGACCTATCCGCCACAGGGTGGAGATCGACATCGATCGCTGGAATCC 311  
 QY 213 IleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySer 232  
 Db 312 ATCGCCGACAGAACACACCCGCTATGTCATCATATAACCCCAACACCCGTCGGGAGC 371  
 QY 233 ValTyrSerTyrAspHisLeuAlaLysValIleAlaGluValAlaArgLysLeuGlyIleLeu 252  
 Db 372 GTTTACTCTACGACCATCTGCAAGGTGCGGAGGTGGGAGGTGGGAGGTGGGAGTATG 431  
 QY 253 ValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMet 272  
 Db 432 GTGATGTGACAGGTATACGGCAAGTGGTCTTGGCAGCGCCCGCTTATCCCAATG 491  
 QY 273 GlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIle 292  
 Db 492 GGAGTGTGTTGGGCACATCACCCCTGCTGTCATAGGCTCTCTGTCACAGTCATGATA 551  
 QY 293 ValProGlyTyrAspLeuGlyTyrValAlaValTyrAspProThrLysIleLeuGluLys 312  
 Db 552 GTGCTGTGATGCGGCTTGGATGGGTAGCGGTACGACCCCAAGATCTTACAGGAA 611  
 QY 313 ThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327  
 Db 612 ACTAAGATCTCTACATCAATTACGAATTACCTCAATGTCGACA 656

RESULT 5  
 CD437584  
 LOCUS  
 DEFINITION  
 ELOIN0502F06.b EndospERM\_5 Zea mays cDNA, mRNA sequence.  
 CD437584  
 VERSION  
 CD437584.1 GI:31353227  
 EST  
 Zea mays  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE  
 1 (bases 1 to 891)  
 AUTHORS  
 Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and  
 Messing, J.  
 TITLE  
 Sequencing of the maize endospERM ESTs  
 JOURNAL  
 Unpublished  
 COMMENT  
 Contact: Lai, Jinsheng  
 Dr. Joachim Messing's lab  
 Waksman Institute, Rutgers University  
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA  
 Tel: 732-445-3801  
 Fax: 732-445-5735  
 Email: jlai@waksman.rutgers.edu  
 Seg primer: T3.

FEATURES  
 source  
 1..891  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="W22"  
 /db\_xref="taxon:4577"  
 /tissue\_type="EndospERM of 7-23DAP"  
 /clone\_lib="EndospERM 5"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI."  
 BASE COUNT 170 a 294 c 259 g 168 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.7e-99 Length: 891  
 Score: 1024.50 Matches: 202  
 Percent Similarity: 84.67% Conservative: 19  
 Best Local Similarity: 77.39% Mismatches: 33

Query Match: 42.87% Indels: 7  
 DB: 14 Gaps: 3  
 US-10-019-783-1 (1-461) x CD437584 (1-891)  
 QY 24 AlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGluTyrAsn 43  
 Db 126 GCGACCCACCCAGCGCAATGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 170  
 QY 44 PheAlaArg---GlyLysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArg 62  
 Db 171 TTCGTCGGCGCGCGAGGAGGTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 230  
 QY 63 AlaIleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeuPro 82  
 Db 231 GCGCGCGCGGTCAAGATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 290  
 QY 83 LeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValAlaGluAsp 102  
 Db 291 CTTGCGCACGGGACCCCTCGTGTTCGGCGCTTCGGCACCGCGCGCGCGCGCGCG 350  
 QY 103 AlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeu 122  
 Db 351 GCGCTGCGCGCGCGCTCCGACGCGCAAGTTCACTGTACCCCGCGCGCGCGCGCGCT 410  
 QY 123 ProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSer 142  
 Db 411 CCGAAGCCCGCGCTGCTCGCGCGAGCACCTGCTCAAGTATCTTCATACAGCTGTCA 470  
 QY 143 AlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProVal 162  
 Db 471 ACTGATGACATCTCTCCCGCTGAGGACCTCAAGCCATTGAGTTGCTGCTCAGTC 530  
 QY 163 LeuAlaGlnThrAlaGlyAlaSerIleLeuLeuProArgProGlyTyrProAsnTyrGlu 182  
 Db 531 CTTGCCCAA---CCGGCGCGCAATATTGCTCCCAAGACCGAGCTATCCAAATATGAG 587  
 QY 183 AlaArgAlaAlaPheAsnLysLeuGluValAlaArgHisPheAspLeuIleProAspLysGly 202  
 Db 588 GCGCGTGGAGGACTGCAACATTCACAGTTCCGCAATTCGATCTGATTCCTGAGAGGG 647  
 QY 203 TrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetVal 222  
 Db 648 TGGGAATTTGATATCGACTCTCTGGAGTGGATGCTGCAAGAACACACCCGCAATGGTC 707  
 QY 223 IleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysVal 242  
 Db 708 ATCATTAACCCCAACAAATCTTGGCGCACTGCTACACCCGCGCAATTTAGCCAGGTC 767  
 QY 243 AlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGlyValTyrGlyLysLeu 262  
 Db 768 GCGAGGTAGCAAGGAAGCTTGAATATCTAGTATCATCGCTGATGAAGTGTATGAAACCTG 827  
 QY 263 ValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeu 282  
 Db 828 GTGTTGGGGACACCCCTTACGTCCTCAATGGGTGCTTTGGCCATATTGCGCTGTGTTG 887  
 QY 283 Ser 283  
 Db 888 AGC 890

RESULT 6  
 CB668395  
 LOCUS  
 DEFINITION  
 OSUNEd16D17.f OSUNEd Oryza sativa (japonica cultivar-group) cDNA  
 clone OSUNEd16D17 5', mRNA sequence.  
 CB668395  
 ACCESSION  
 CB668395.1 GI:29672120  
 VERSION  
 EST.  
 KEYWORDS  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;



EST.  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Erihartoideae; Oryzaceae; Oryza.  
1 (bases 1 to 878)  
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: gta aaa cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 04 row: P column: 10  
Seq primer: gca aac agc tat gac cat g.  
Location/Qualifiers  
1..878  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNEB04P10"  
/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNEB"  
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:  
XhoI; 24 hrs after inoculation with Rice Blast (Che 86061  
)"  
BASE COUNT 262 a 198 c 151 g 267 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,07e-99 Length: 878  
Score: 1020.00 Matches: 188  
Percent Similarity: 89.96% Conservative: 27  
Best Local Similarity: 78.66% Mismatches: 24  
Query Match: 42.68% Indels: 0  
DB: 14 Gaps: 0  
US-10-019-783-1 (1-461) x CB643965 (1-878)  
QY 221 MetValIleLeuAsnProCysGlySerValTyrSerValTyrAspHisLeuAla 240  
Db ATAGTCATCATCAATCCCAATATCCATCGGGAATGTGTACATTAGGACATTATCC 819  
QY 241 LysValAlaGluValAlaArgLysLeuGlyLysLeuValIleAlaAspGluValTyrGly 260  
Db AAGGTGGCAGAGGTAGCAAGGAAGCTTGGGATATTGCTAATCTGATGAGGTGATGGT 759  
QY 261 LysLeuValLeuGlySerAlaProPheLeuProMetGlyValPheGlyHisIleAlaPro 280  
Db AATTGGTTTGGGAGTTCCTCCCAATGGTTCCTTGGGACATCGTACCA 699  
QY 281 ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTyr 300  
Db ATATTACCATAGGATCGCTATCAAGAGGTGGTAGTCCCGGATGGGACATGGTTGG 639  
QY 301 ValAlaValTyrAspProThrLysIleLeuGlyLysLeuValIleAlaAspGluValTyr 320  
Db GTAGCAATATGTGACCCCAAGAGCTCTACAGAAACCAAGATTGCAACATTAACTACT 544  
QY 321 AsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIle 340  
Db AATTTCCTTAATGTTTCAACTGATCCAGCACTTTCATTACGGGAGCTCTACCGAATATT 484  
QY 341 LeuGluAsnThrIleValAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSer 360  
Db CTTAAGAAATCCAGGAAGAACTTTTAAGAGGATAATGTATTTGCTTACGGAAACATCA 424  
QY 361 GluIleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGlu 380  
Db GATATTTCGTATAGAGGATAAAGGATATTAAATGCATCACTTGTCTCTCAACACCCGAA 364  
QY 381 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAsp 400  
Db GGATCATGTTTGTGATGTGAATGAACCTATATCTTTTGGAGGGAATCCATGATGAT 304  
QY 401 IleAspPheCysCysLysLeuAlaLysGluLysSerValIleLeuCysProGlySerVal 420  
Db GTTGATTTTGTGCAACTTCGGAAGAGAGTGGTGATCTTTGCCAGGAGGTGTG 244  
QY 421 LeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 440  
Db CTGGGAATCAAGAATTTGGTTCGCATTACTTTTGTATTTGATTTCTCTCTCTCTGGAT 184  
QY 441 GlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysLysAsnSerIleAsn 459  
Db GGTCTTGAGAGGATCAATCTCTTCTGCCAAGGCACAGAAAGAAACCCCTTNAAT 127  
RESULT 8  
CB643965/c  
LOCUS  
DEFINITION  
OSJNEB04P10.r OSJNEB Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNEB04P10 3', mRNA sequence.  
CB643965  
ACCESSION  
CB643965.1 GI:29638956

```

Db 578 AATTTCCTTAATGTTCAACTGATCCAGCACTTTCATTGAGGAGCTTACCGCAATATT 519
Qy 341 LeuGluAsnThrLysAlaAspPhePheLysArgIleLeuGlyLeuLysGluSerSer 360
Db 518 CTTAAGATACACAGGAAGAATCTTTAAGAGGATTAATGTTTCTTACCGAAACATCA 459
Qy 361 GluLeuCysTyrArgGluLeuLysGluAsnLysTyrIleThrCysProHisLysProGlu 380
Db 458 GATATTTCCTATAGAGGAATAAAGGATTAATGATCACTTCTCCTCAAGCCGAA 399
Qy 381 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluLeuHisAsp 400
Db 398 GGATCCATGTTTGATGGTGAATTAATGACCTATATCTTTTGGAGGAAATCCATGAT 339
Qy 401 IleAspPheCysLysLeuAlaLysGluGluSerValLleLeuCysProGlySerVal 420
Db 338 GTTGATTTTGTGCAACTTGGCAAGAGAGTCGGTGAATCTTTTGGCAGGAGGTGTG 279
Qy 421 LeuGlyMetGluAsnThrValArgIleThrPheAlaCysValProSerSerLeuGluAsp 440
Db 278 CTGGGAATCAAGAAATGGTTCGATTAATCTTTTCTTATGATTCATCTCTCTCGAT 219
Qy 441 GlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysLysAsnSerIleAsn 459
Db 218 GGTCTTGAGAGATCAATCTCTCTGCCAAGGCACAGAGAAACCCCTTTAAAT 162

```

```

RESULT 9
CB618033/c
LOCUS
DEFINITION
OSJNE041J23.r OSJNEa Oryza sativa (indica cultivar-group) cDNA
clone OSJNEa01J23 3', mRNA sequence.
ACCESSION
CB618033.1
VERSION
CB618033.1
KEYWORDS
ORIGIN
Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 902)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: J column: 23
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1..902
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSJNEa01J23"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEa"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Lesion Mimic SPL 11"
XhoI 270 a 202 c 153 g 277 t

```

## FEATURES

source

```

ORIGIN
Alignment Scores:
Pred. No.: 5,28e-99 Length: 902
Score: 1020.00 Matches: 188
Percent Similarity: 89.9% Conservative: 27
Best Local Similarity: 78.6% Mismatches: 24
Query Match: 42.6% Indels: 0
DB: 14 Gaps: 0
US-10-019-783-1 (1-461) x CB618033 (1-902)
Qy 221 MetValIleLeuAsnProCysGlySerValTyrSerTyrAspHisLeuAla 240
Db 901 ATAGTCATCATTAATCCCAATTAATCCATCGGGATGTGTACACTTACGAGCAATTAACC 842
Qy 241 LysValAlaGluValAlaArgLysLeuGlyLleLeuValIleAlaAspGluValTyrGly 260
Db 841 AAGGTGGCAGAGGTAGCAAGGAAGCTTGGGATATTGGTAATTACTGTAGGTGTATGTT 782
Qy 261 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro 280
Db 781 AATTGTTTTTGGGAGTTCCCAATTTGTTCCCAATGGGTGCTTTGGGCACATCTACCA 722
Qy 281 ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrp 300
Db 721 ATATTAACCATAGGATCGCTATCAAGAGGTGGATAGTCCGGGATCGGACTTGGTTGG 662
Qy 301 ValAlaValTyrAspProThrLysLysLeuGluLysThrLysLysSerThrSerIleThr 320
Db 661 GTAGCAATATGTGACCCCAAGAGACTCTACAGAAACCAAGATTGCAACATTAATTAAT 602
Qy 321 AsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIle 340
Db 601 AATTTCCTTAATGTTCACTGATCCAGCAACTTTCAATTCAGGAGCTCTACCGAATATT 542
Qy 341 LeuGluAsnThrLysAlaAspPhePheLysArgIleLeuGlyLeuLysGluSerSer 360
Db 541 CTTAAGATACCAAGGAAGAATCTTTAAGAGGATAATTGATTTCCTTACGGAACATCA 482
Qy 361 GluLeuCysTyrArgGluLysGluLeuLysLysTyrIleThrCysProHisLysProGlu 380
Db 481 GATATTGCTATAGAGGATAAAGGATTAATTAATGCACTTCCTTCTCAGAGCCGAA 422
Qy 381 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluLleHisAsp 400
Db 421 GGATCCATGTTTGTGATGGTGAATTAATGAACTATATCTTTTGGAGGAAATCCATGAT 362
Qy 401 IleAspPheCysLysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerVal 420
Db 361 GTTGATTTTGTGTCCAACTTGGAAAGAGAGTGGTGATTTCTTCCAGGAGGTGTG 302
Qy 421 LeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGluAsp 440
Db 301 CTGGGAATCAAGAAATGGGTTCGCATTACTTTTGTCTATTGATTCATCTCTCTCGAT 242
Qy 441 GlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysLysSerIleAsn 459
Db 241 GGTCTTGAGAGGATCAATCTCTTCTGCCAAGGCACAGAGAAACCCCTTTAAAT 185

```

## RESULT 10

```

CB618033/c
LOCUS
DEFINITION
OSJNE04F21.r OSJNEe Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEe04F21 3', mRNA sequence.
ACCESSION
CB618033
VERSION
CB618033.1
KEYWORDS
ORIGIN
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
849 bp mRNA linear EST 09-APR-2003

```

BASE COUNT

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 849)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088 USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 04 row: F column: 21
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1..849
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNE04F21"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNE"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
BASE COUNT 260 a 190 c 140 g 259 t
ORIGIN
Alignment Scores:
Pred. No.: 6,18e-99 Length: 849
Score: 1019.00 Matches: 188
Percent Similarity: 89.92% Conservative: 26
Best Local Similarity: 78.99% Mismatches: 24
Query Match: 42.64% Indels: 0
Gaps: 0
14
US-10-019-783-1 (1-461) x CB671005 (1-849)
QY 222 ValIleLeuAsnProAsnAsnProCysGlySerValTySerTyAspHisLeuAlaLys 241
Db 847 STCATATAAATCCCAATAATCCATCGGGAATGTGTACACTTACGACATTTATCCAAG 788
QY 242 ValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyGlyLys 261
Db 787 GTGGCAGAGGTAGCAAGGAGCTTGGGATATTGGTAATTACTGATGAGGTGTAATGGTAAT 728
QY 262 LeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaProVal 281
Db 727 TTGTTTGTGGAGTTCCTCCATTTGTCCTCCATGGTTCCTTGGCACATCGTACCAATA 668
QY 282 LeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpVal 301
Db 667 TTACCATAGATCGCTATCAAGAGGTGGATAGTGGCGGATGCGGACTTGGTGGGTA 608
QY 302 AlaValTyAspProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsn 321
Db 607 GCAATATGTGACCCCAAGACACTCTACAAGAACCAAGATTCGAACATTAATTAATTAAT 548
QY 322 TyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeu 341
Db 547 TTCTTAATGTTTCAACTGATCCGCACTTTCATTTCAGGGAGCTCTACCGAATATCTT 488
QY 342 GluAsnThrLysAlaAspPheLysArgIleIleGlyLeuLeuLysGluSerSerGlu 361
Db 487 AAGATACCAAGGAGAGATTTCTTTAAGAGGATAATTGATTGCTTACGGAACATCAGAT 428

```

```

QY 362 IleCysTyArgGluIleLysGluAsnLysTyIleThrCysProHisLysProGluGly 381
Db 427 ATTTGCTATAGAGGATAAAGGATTAATAATGTCATCACTTCTCTCAAGCCGGAAGGA 368
QY 382 SerMetPheValMetValLysLeuAsnLeuHisLeuGluGluIleHisAspAspIle 401
Db 367 TCCATGTTTGATGGTGAATTCACCTATATCTTTGGAGGGAATCCCATGATGATGT 308
QY 402 AspPheCysCysLysLeuAlaLysGluLysSerValIleLeuCysProGlySerValLeu 421
Db 307 GATTTTGTGCCAACTTGGGAAAGAGTCGGTGATTCCTTCCCCAGGAGTGCTG 248
QY 422 GlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGly 441
Db 247 GGAATGAGGAATTCGGTTCGCAATTCATTTTGCATTATTCATCTTCTCTCTGATG 188
QY 442 LeuGluArgValLysSerPheCysGlnArgAsnLysLysLysAsnSerIleAsn 459
Db 187 CTTGAGAGGATCAATCTTCTGCAAAGGCACACAGAGAAACCCCTTTAAT 134
RESULT 11
CB652840/c
LOCUS CB652840 832 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEC03E07.r OSJNEC Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEC03E07 3', mRNA sequence.
ACCESSION CB652840
VERSION CB652840.1 GI:29656565
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 832)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 03 row: E column: 07
Seq primer: gga aac agc tat gac cat g.
FEATURES
Location/Qualifiers
1..832
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC03E07"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEC"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT 256 a 186 c 140 g 250 t
ORIGIN
Alignment Scores:
Pred. No.: 1.62e-98 Length: 832
Score: 1015.00 Matches: 187
Percent Similarity: 89.87% Conservative: 26
Best Local Similarity: 78.90% Mismatches: 24

```

Query Match: 42.47% Indels: 0  
DB: 14 Gaps: 0

US-10-019-783-1 (1-461) x CB652840 (1-832)

QY 223 IleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysVal 242  
DB 831 ATCATAAATCCATAATCCATGGGGAATGTACATTCAGAGCATTTATCCAGGTG 772  
QY 243 AlaGluValAlaArgLysLeuGlyLeuValLleAlaAspGluValTyrGlyLysLeu 262  
DB 771 GCAGAGGTAGCAAGCAAGCTGGGATATGGTAATTAAGTATGAGGTGTATGTAATTTG 712  
QY 263 ValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeu 282  
DB 711 GTTTTGGGAGTCCCAATTTGTCCTAATGGGTGCTTTGGGCACATCGTACCAATATTA 652  
QY 283 SerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAla 302  
DB 651 ACCATAGGATCGTATCAAGAGGTGGATAGTGGCGGATGGCGACTTGGTTGGGTAGCA 592  
QY 303 ValTyrAspProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyr 322  
DB 591 ATATGTGCCCCCAAGAACACTACAGAAACCAAGATTGCAACATTAATTAATTAATTTT 532  
QY 323 LeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGlu 342  
DB 531 CTATATGTTTCACTGATCCAGCACTTTCATTCAGGAGGCTTACCGAATATTTCTTAAG 472  
QY 343 AsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIle 362  
DB 471 AATACCAAGGAAGAAATCTTTAAGAGGATAATGTTTGTTCAGGAAACATCAGATATT 412  
QY 363 CysTyrArgGluIleLysLysLysLysLysLysLysLysLysLysLysLysLysLys 382  
DB 411 TCTTATAGGAGTAATAGGATATTAATGCATCATCTTCTCTCAAGCCCGAGGATCC 352  
QY 383 MetPheValMetValLysLeuAsnLeuHisLeuLeuGluIleHisAspIleAsp 402  
DB 351 ATGTTTGTGATCGTGAATATGAACCTATATCTTTTGGAGGGAATCCATGATGTTGAT 292  
QY 403 PheCysCysLysLeuAlaLysGluSerValIleLeuCysProGlySerValLeuGly 422  
DB 291 TTTTGTGCGCACTTGCAGAAAGAGTGGTGATCTTTGCCCGAGGAGTGTCTGGGA 232  
QY 423 MetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeu 442  
DB 231 ATGAAGATTGGTTCGCATTACTTTTGTATTGATTTCATCTCTCTCTGATGTTCTT 172  
QY 443 GluArgValLysSerPheCysGlnArgAsnLysLysLysAsnSerIleAsn 459  
DB 171 GAGAGATCAATATCTTCTGCCAAGGACACAGAAAGAAAACCCCTTTAAAT 121

## RESULT 12

## CB646007/c

## LOCUS

## DEFINITION

## OSUNEB08C06.r OSUNEB Oryza sativa (japonica cultivar-group) cDNA

## clone OSUNEB08C06 3', mRNA sequence.

## CB646007

## CB646007.1 GI:29641000

## EST.

## Oryza sativa (japonica cultivar-group)

## Oryza sativa (japonica cultivar-group)

## Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

## Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

## Ehrhartoideae; Oryzaceae; Oryza.

## 1 (bases 1 to 883)

## Jantasurivarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,

## Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.

## Large-scale identification of ESTs involved in the interaction

## between rice and Magnaporthe grisea

## Unpublished

## Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0086, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: <http://genome.arizona.edu>

PCR Primers

FORWARD: gta aaa cga cgg cca gtc

BACKWARD: gga aac agc tat gac cat g

Place: 08 row: C column: 06

Seq primer: gga aac agc tat gac cat g.

Location/Qualifiers

1..883

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="OSUNEB08C06"

/tissue\_type="Leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_lib="OSUNEB"

/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:

XhoI; 24 hrs after inoculation with Rice Blast (Che 86061

)"

BASE COUNT 265 a 199 c 151 g 268 t

ORIGIN

Alignment Scores:

Pred. No.: 5,68e-97 Length: 883

Score: 1001.00 Matches: 184

Percent Similarity: 89.41% Conservatives: 27

Best Local Similarity: 77.97% Mismatches: 25

Query Match: 41.88% Indels: 0

DB: 14 Gaps: 0

US-10-019-783-1 (1-461) x CB646007 (1-883)

QY 224 IleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAla 243

DB 881 ATAAATCCCAATAATCCATCGGGAATGTACATTCAGAGCATTTATCCAGGTGGGCA 822

QY 244 GluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuVal 263

DB 821 GAGTGAGCAGGAGGAGTGGGATATGGTAATTAATGATGAGGTGTATGTAATTTGTT 762

QY 264 LeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSer 283

DB 761 TTTGGGAGTCCCAATTTGTCCTAATGGGTGCTTTGGGCACATCGTACCAATATTAACC 702

QY 284 IleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaVal 303

DB 701 ATAGGATCGCTATCAAGAGGTGGATAGTCCCGGATGGGACTTGGTTGGGTAGCAATA 642

QY 304 TyrAspProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeu 323

DB 641 TGTGACCCCAAGAAGACTCTACAGAAACCAAGATTGCAACATTAATTAATTAATTCCTT 582

QY 324 AsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsn 343

DB 581 AATGTTTCAACTGATCCAGCACTTTTCATTTCAGGAGGAGTCTACCGAATATTTCTTAAGAT 522

QY 344 ThrLysAlaAspPhePheLysArgIleIleGlyLeuLysGluSerSerGluIleCys 363

DB 521 ACCAAGGAAGAAATCTTTAAGAGGATAATTTGTTCTTACGAAACATCAGATATTTGC 462

QY 364 TyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMet 383

DB 461 TATAGAGGATAAAGGATATTAATGCATCTTGTCTTCAAGCCCGAGGATCCATG 402

QY 384 PheValMetValLysLeuAsnLeuHisLeuLeuGluIleHisAspIleAspPhe 403



```

Db      401 TTGTGATGGTGAATGAACCTATATCTTTTGGAGGGAATCCATGATGATTTGATTTT 342
Qy      404 CysCysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMet 423
Db      341 TGTTCACAACTTCGCAAGAGAGTGGTGATTCITTTGCCAGGAGTGTGCTGGGATG 282
Qy      424 GluAsnTrpValArgIleThrPheAlaCysValProSerSerIleGlnAspGlyLeuGlu 443
Db      281 AGAATATGGTTCGCAATTAATTTTGTATTAATCATCTTCTCTCTGATGCTTGTAG 222
Qy      444 ArgValLysSerPheCysGlnArgAsnLysLysAsnSerIleAsn 459
Db      221 AGGATCAATCTTCTGCCAAGGACACAGAAAGAAACCTTTAAAT 174

RESULT 13
CB618032
LOCUS   OSIIeA01J23.f OSIIeA Oryza sativa (indica cultivar-group) cDNA
DEFINITION
ACCESSION   CB618032
VERSION     CB618032.1 GI:29613019
KEYWORDS    EST.
SOURCE      Oryza sativa (indica cultivar-group)
ORGANISM    Oryza sativa (indica cultivar-group)
REFERENCE   1 (bases 1 to 809)
AUTHORS     Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
            Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE       Large-scale identification of ESTs involved in the interaction
            between rice and Magnaporthe grisea
JOURNAL     Unpublished
COMMENT     Contact: Rod Wing
            Arizona Genomics Institute
            University of Arizona
            Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
            85721-0088, USA
            Tel: 520 626 3967
            Fax: 520 621 9288
            Email: http://genome.arizona.edu
            PCR Primers
            FORWARD: gta aaa cga cgg cca gtg
            BACKWARD: gga aac agc tat gac cat g
            Plate: 01 row: J column: 23
            Seq primer: gta aaa cga cgg cca gtg.
            Location/Qualifiers
            1..809
            /organism="Oryza sativa (indica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="IR36"
            /db_xref="taxon:39946"
            /clone="OSIIeA01J23"
            /tissue_type="Leaf"
            /dev_stage="3 week"
            /lab_host="DH10B"
            /clone_lib="OSIIeA"
            /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
            XhoI; Lesion Mimic SPL 11"
BASE COUNT 176 a 245 c 235 g 153 t
ORIGIN

Alignment Scores:
Pred. No.: 4.09e-96 Length: 809
Score: 992.50 Matches: 200
Percent Similarity: 82.08% Conservative: 29
Best Local Similarity: 71.68% Mismatches: 34
Query Match: 41.53% Indels: 16
DB: 14 Gaps: 6

US-10-019-783-1 (1-461) x CB618032 (1-809)
Qy      2 ValHisGln-----SerAsnGlyHisGly-GluAlaAlaAlaAlaAsnG1 18

```

```

Db      6 ATTCAACAAACACAAAAACAGAGCAGCGCCATGGCACCCAGCGCGCGCGCGCG---GC 62
Qy      18 YLysSerAsnGlyHisAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAla 38
Db      63 GAGCAGCAACGCGC-----GGCGCGGAGAGCGCGCGC-----AG 95
Qy      38 aAlaValGluTrpAsnPhe---AlaArgGlyLysAspGlyLysLeuAlaThrThrGlyAl 57
Db      96 CAGCAAGGAGTGGAGGCTGACGGCGCGACGAGGGGCGCGCGATGGCGCGCGCGGGGA 155
Qy      57 aLysAsnSerIleArgAlaIleArgTyLysIleSerAlaSerValGluGluSerGlyPr 77
Db      156 CAAGATGAGCATCGGCGCGTCCGGTACAAAGATCAGCGCGCAGCGCTGCACACCGCGCC 215
Qy      77 oArgProValLeuProLeuAlaHisGlyAspProSerValPheProSerValPheArgThrAl 97
Db      216 GCGCCCGCTCTCGCGCTGCCCGGACCGCCCTCGTGTTCCTCCGAGTTCGCGACCGC 275
Qy      97 aValGluAlaGluAspAlaValAlaAlaLeuAlaArgThrGlyGlnPheAsnCysTyAl 117
Db      276 CGCGGAGCGGAGGACCGCTCGCGCGCGCTCCGCTCGCGCGACTTCAACTGCTACCC 335
Qy      117 aAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyVa 137
Db      336 CGCGCGCTCGGCTCCCGCGCGGACGCTGTGTGGCAGATCATTTGTACGCGACCT 395
Qy      137 lProTyLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGl 157
Db      396 CCCATACAGCTATCTTCGATGACATCTTCCTAACCGCTGGAGGAACTCAGGCCATCGA 455
Qy      157 uValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGl 177
Db      456 GGTGCTGAATCTCAATCTTCCCAA---CCTGGCACAAACATATTGCTTCCTAGACAGG 512
Qy      177 YTyProAsnTyTrpGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLe 197
Db      513 CTACCCAAACTATGAAGCTCGAGCGCGCTTCAACACCTTGAAGTTGCTACCTTGAATCT 572
Qy      197 uIleProAspLysGlyTrpGluIleAspLeAspSerLeuGluSerIleAlaAspLysAs 217
Db      573 TATTCCTGAGAGGCTGGGAGATTGACCTTAACCTCCCTAGAAATCTATTGGCGACAAGAA 632
Qy      217 nThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTySerTyAs 237
Db      633 CACTACTGGATAGTCAATATAATCCCAATATCCATTCATCGCGGAATGTGTACACTTACGA 692
Qy      237 pHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGl 257
Db      693 GCATTTATCCAAAGTGGCAGAGTAGCAAGGAAGCTTGGATATTGGTAATTACTGATGA 752
Qy      257 uValTyGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPhe 275
Db      753 GGTGATGTAATTTGGTTTTTGGGAGTTCCCAATTTGTTCCCAATGGGTGCTTT 807

```

```

RESULT 14
EG313019/c
LOCUS   EG313019
DEFINITION   Wheat endosperm cDNA library Triticum aestivum
            cDNA clone WHE0091_D01_G01, mRNA sequence.
ACCESSION   EG313019
VERSION     EG313019.1 GI:13114822
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
ORGANISM    Triticum aestivum

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 619)
AUTHORS     Altenbach,S., Anderson,O.D., Chao,S., Galili,G., Han,P.S., Hsia
            ,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and
            Tong,J.C.

```

The structure and function of the expressed portion of the wheat

Qy	60	Ser-IleArgAlaIleArgTyrIlyssIleSerAlaIleValGluGluSerClyIProArgPro	79
Db	11	AGCATCCGGCGGATACGGTACAGATACACGCGAGCGTGCAGGAAGGGCGCGGGCC	70
Qv	80	ValLeuProLeuAlaHisIcglVasPProSerValPheProAlaPheArgTyrAlaValGlu	99

```
Db      71  GTGCTGCGCGTGGCCACGCGGACCGCTCGGTGTTCCCGGCTTCGCGACGCGCGTGGAG 130
QY      100  AlaGluAspAlaValAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGly 119
Db      131  GCGAGGACGCGTCCCGCGTGGCGACCGCGAGTTCAACTGTCTACCCCGCGGC 190
QY      120  ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerClnGlyValProTyr 139
Db      191  GTCGGCTTCCCGCGCCACGAGAGCGCGTGGGAGAGACCTGTGCGAGGCGGTGCCGTAC 250
QY      140  LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159
Db      251  ATGCTATCGCGCGACGACGCTCTCTCCACCGCGCGGACCCAGCGATCGAGGTGATA 310
QY      160  IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 179
Db      311  ATCCCGGTCTGGCCCGAGACCCCGCGCCCAACATTCTGCTCCCGAGCCAGGCTACCCA 370
QY      180  AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199
Db      371  AACTACGAGCGCGCGCGCTCAACAGGCTGGAGTCCGGCATTTGACCTCATCCCC 430
QY      200  AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219
Db      431  GACAAAGGGTGGAGATCGACATCGACTCGCTGGAATCCATCGCGCACAGAACACCACC 490
QY      220  AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239
Db      491  GCCATGGTTCATATAAACCCCAACACCGTGGCGCAGCGTTTACTCTACGACCATCTG 550
QY      240  AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGlu 257
Db      551  TCCAAAGGTGCGGAGGTGGCGAAAGGCTCGGAATATTGGTCATTGCTGACGAG 604
```

Search completed: November 9, 2003, 01:19:51  
Job time : 2826.28 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 17:15:50 ; Search time 5078.74 Seconds  
(without alignments)  
3713.386 Million cell updates/sec

Title: US-10-019-783-1

Perfect score: 2390

Sequence: 1 MVHQNNGHGEAAAAANGKS.....LKVKSFCQRNKKNSINGC 461

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2888711 seds, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10019783/runat\_07112003\_161643\_19836/app\_query.fasta\_1.1358  
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=ptc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORX=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10019783 @CGN 1.1 6944 @runat 07112003\_161643\_19836 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.em.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rat.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sv.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2390	100.0	1660	8	D88273	D88273 Hordeum vul
2	2142.5	89.6	1895	8	AB005788	AB005788 Hordeum v
3	1988	83.2	10966	6	BD012010	BD012010 Creation
4	1988	83.2	10966	6	E49893	E49893 Creation of
5	1988	83.2	10966	8	AB024006	AB024006 Hordeum v
6	1228.5	51.4	1290	6	AX653727	AX653727 Sequence
7	1150.5	48.1	920	6	AX653224	AX653224 Sequence
8	1128	47.2	1389	6	AX575731	AX575731 Sequence
9	1127	47.2	1600	8	AY054204	AY054204 Arabidops
10	1127	47.2	1611	8	AF301900	AF301900 Arabidops
11	1125	47.1	1611	8	AF301899	AF301899 Arabidops
12	1124	47.0	1389	6	AX507289	AX507289 Sequence
13	1124	47.0	1420	8	AY051293	AY051293 Arabidops
14	1124	47.0	1611	8	AF301898	AF301898 Arabidops
15	1124	47.0	1683	8	AY050987	AY050987 Arabidops
16	1122	46.9	1243	6	AX575733	AX575733 Sequence
17	1122	46.9	1276	8	BT001912	BT001912 Arabidops
18	1121	46.9	1570	8	BT000782	BT000782 Arabidops
19	1120	46.9	1333	6	AX654009	AX654009 Sequence
20	1110.5	46.5	1236	8	CBL458993	CBL458993 Coleus bl
21	1100	46.0	1564	8	AY113848	AY113848 Arabidops
22	1023.5	42.8	1338	6	AX412376	AX412376 Sequence
23	1023.5	42.8	1338	6	AX507174	AX507174 Sequence
24	1023.5	42.8	1338	6	AX651743	AX651743 Sequence
25	1023.5	42.8	1338	8	BT006593	BT006593 Arabidops
26	1023.5	42.8	1495	8	BT002475	BT002475 Arabidops
27	1017.5	42.6	1535	8	AY085324	AY085324 Arabidops
28	1012.5	42.4	1334	6	AX575729	AX575729 Sequence
29	1003.5	42.0	126315	2	AP004116	AP004116 Oryza sat
30	1002	41.9	159980	2	AP005743	AP005743 Oryza sat
31	953	39.9	93695	8	ATF2009	ATF2009
32	953	39.9	199749	8	ATCHRIV68	ATCHRIV68
33	928.5	38.8	70475	8	AB018112	AB018112 Arabidops
34	903	37.8	61394	8	AC007048	AC007048 Arabidops
35	892.5	37.3	73840	8	AB007644	AB007644 Arabidops
36	881	36.9	151073	2	AC134624	AC134624 Oryza sat
37	878	36.7	1201	8	AY142527	AY142527 Arabidops
38	872	36.5	1634	8	AY099811	AY099811 Arabidops
39	871.5	36.5	930	6	AX653614	AX653614 Sequence
40	871.5	36.5	1269	6	AX575727	AX575727 Sequence
41	871.5	36.5	1462	8	AF268090	AF268090 Arabidops
42	871.5	36.5	1495	8	BT000307	BT000307 Arabidops
43	861.5	36.0	1284	8	AY187882	AY187882 Brassica
44	849	35.5	1306	8	AY123024	AY123024 Arabidops
45	849	35.5	1563	8	AY070389	AY070389 Arabidops

# ALIGNMENTS

RESULT 1

[illegible]

```

QY 401 IleAspPheCysCysLysLeuAlaLysGluSerValIleLeuCysProGlySerVal 420
Db 1262 ATAGATTTTGTGCGAGCTCCAAAGAGAAATCAGTAAATTTATGTCAGGGAGTGT 1321
QY 421 LeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 440
Db 1322 CTTGGAATGAAATTTGGTGGCTATTACTTTTGGCTGCTCCATCTCTCTCAAGAT 1381
QY 441 GlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysLysAsnSerIleAsnGly 460
Db 1382 GACATCGAAGGGTCAATCATCTCTCAAGGAAAGAAAGAAATTCATAATGGT 1441
QY 461 Cys 461
Db 1442 TGT 1444

RESULT 2
AB005788
LOCUS
DEFINITION
Hordeum vulgare mRNA for nicotianamine aminotransferase B, complete
cds.
ACCESSION
AB005788
VERSION
AB005788.1 GI:6469086
KEYWORDS
nicotianamine aminotransferase B.
SOURCE
Hordeum vulgare subsp. vulgare
ORGANISM
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
REFERENCE
1 (sites)
Takahashi,M., Yamaguchi,H., Nakanishi,H., Shioiri,T.,
Nishizawa,M.K. and Mori,S.
Cloning two genes for nicotianamine aminotransferase, a critical
enzyme in iron acquisition (Strategy II) in graminaceous plants
Plant Physiol. 121 (3), 947-956 (1999)
20027437
PUBMED
10557244
REFERENCE
2 (bases 1 to 1895)
Mori,S.
Direct Submission
Submitted (08-JUL-1997) Satoshi Mori, The University of Tokyo,
Plant Molecular Physiology; Yayoi 1-1-1, Bunkyo-ku, Tokyo 113,
Japan (E-mail: aal07@hongo.ecc.u-tokyo.ac.jp,
Tel:03-3812-2111(ex.5106), Fax:03-3812-0544),
Location/Qualifiers
1..1895
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Ehimedakaka No.1"
/db_xref="taxon:112509"
/tissue_type="root"
1..1895
/gene="naat-b"
76..1731
/gene="naat-b"
/feature="naat-b"
/codon_start=1
/product="nicotianamine aminotransferase B"
/protein_id="BA087053.1"
/db_xref="GI:6469087"
/translation="MATVROS DGVAANG LAVALAAANGKSNHGVAANGKSNHGVD
ADANGKNGHVAADANGKNGHAEATANGHAGATANGKNGHRENSNGHAEADANGE
SNEHAEBSAANGESNGHAAAEAEAEVWENFAGKDGVLAAATGAMNSIRAIKYKISA
SVQFKGPRPVLPLAHGDPSPVPAPRTAVEAEADAAMRTGQNCIPAGVGLPAARSA
VAEHLSGVPMLSADVDVFTAGTQAEVILIPVLAQTAGANILLPRPGVYFEARAA
FNRLEVRHFLIDPKWEIDIDLESADXTNTAMVINENPNPCGSWYSDHLSKVAE
VAKRLGLVTADEVKGLVDSAPFIPMGVFGHTPVLISGLSKSWIVPGWRLGWA
VYDFPKLQTKTKSTITNYLNTSDPATFIQAALPOLIENTKEDFFKAIIGLKES
EICVQIKENKTYTCPKPEGSFMVVKLNHLLEIDDDIDDFCCCKLAKESVILCPG
SVLGMANVWRITACVPSISQDGLGRKISKFCQKKNKSNSSDDC"
1895
/gene="naat-b"
/feature="15 A nucleotides"
polya_site

```

```

BASE COUNT 447 a 534 c 546 g 368 t
ORIGIN
Alignment Scores:
Pred. No.: 1,74e-144 Length: 1895
Score: 2142.50 Matches: 415
Percent Similarity: 93.55% Conservative: 20
Best Local Similarity: 89.25% Mismatches: 21
Query Match: 89.64% Indels: 9
DB: 3
US-10-019-783-1 (1-461) x AB005788 (1-1895)

```

```

QY 4 GlnSerAsnGlyHisGlyGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaHis 23
Db 340 GAGAGCAACGGCCATGCTGAG-----GCCGCCGACGCCAACGCCGAGACAACGAGCAT 393
QY 24 Ala-----AlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAla----- 38
Db 394 GCCGAGGACTCCCGCGGCGGACGCGGAGAACGCGCATCCGCGCGGCGGCGGAGAGGAG 453
QY 39 -----AlaValGluTrpAsnPheAlaArgGlyLysAspGlyIleLeuAlaThrGly 56
Db 454 GAGGAGCGGTGGAGTGAATTTCCGGGGTGCCACGACCGCGTGTGGCGGCGACGGGG 513
QY 57 AlAlAsnSerIleArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGly 76
Db 514 GCACATGAGCATCCGGCGGATACGATACAGATCAGCGCGAGCGGTGAGGAGAGAGGGG 573
QY 77 ProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThr 96
Db 574 CCGCGCGCGGTGTGCGGTGGCCACGCGGACCGCGTGTCCGGGCTTCCGACG 633
QY 97 AlaValGluAlaGluAspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnGlyTyr 116
Db 634 GCCGTGAGCGCGGAGACCGCGTGGCGCGCGGTGCGCACCGCGCAGTCACTGCTAC 693
QY 117 AlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGly 136
Db 694 CCGCGCGCGGTGGCGGTCCCGCGCGCACGAGGCGCGTGGCGAGACCTGTCGCGAGGC 753
QY 137 ValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIle 156
Db 754 GTCCGTACATGCTATCGCGCGACGAGTCTTCTTCCACCGCGCGGCGGAGGAGGATC 813
QY 157 GluValIleLeuProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgPro 176
Db 814 GAGTGTATATCCCGGTGTGTGCCCGAGACCGCGCGCGCGCAACATTCTGCTCCCGAGCCA 873
QY 177 GlyTyrProAsnTyrGluAlaAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAsp 196
Db 874 GGCTACCCAACTACGAGCGCGCGCGGTTCACAGGCTGAGGTCGCGCATTTCCGAC 933
QY 197 LeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLys 216
Db 934 CTCATCCCGCAAGGGGTGGAGATCGACATCGACTCGCTGGAATTCATCCCGCACAG 993
QY 217 AsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyr 236
Db 994 AACACCGCGCCATGGTGTATATAACCCCAACACCGGTGGCGGAGCGTTTACTCTTAC 1053
QY 237 AspHisLeuAlaLysValaGluValaAlaArgLysLeuGlyIleLeuValIleAlaAsp 256
Db 1054 GACCATCTGTCCAAAGTGGCGGAGGTGGCGAAGAGCTCGGAATATTGGTATTGCTGAC 1113
QY 257 GluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly 276
Db 1114 GAGGTATACGCAAGCTGGTCTGGCGAGCGCGCGTTCATCCCAATGGGAGTGTGGG 1173
QY 277 HisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyr 296
Db 1174 CACATCACCCCTGTGTCTCCATAGGGTCTCTGTCCAAGTATGATGATGCTGATGG 1233

```

Qy	297	ArgLeuGlyTrpValAlaValTyAspProThrLysIleLeuGluLysThrLysIleSer	316
Db	1234	CGGCTTGGATGGGTAGCGGTGTACGACCCAGAAAGATCTTACGAGAACTAAGATCTCT	1293
Qy	317	ThrSerIleThrAsnTyIleuAsnValSerThrAspProAlaThrPheValGlnGluAla	336
Db	1294	ACATCAATTACGAATATACCTCAATGTCGCACAGACCCAGCAACCTTCATTTCAGGCAGCT	1353
Qy	337	LeuProCysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIleLeuLeu	356
Db	1354	CTTCCTCAGATTCTTGAGAACACAAAGGAAGATTTCTTTAAGCGGATTAATGGTCTGCTA	1413
Qy	357	LysGluSerSerGluIleCysTyArgGluIleLysGluAsnLysTyIleThrCysPro	376
Db	1414	ARGGAATCATCAGAGATATGCTCAACAATAAAGGAAAAACAATACATTACATGTCCT	1473
Qy	377	HisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGlu	396
Db	1474	CACAGCCAGAGGATCAATGTTGTTCATGGTGAACCTGAACCTTACATCTTTTGGAGAA	1533
Qy	397	IleHisAspAspIleAspPheCysCysLysIleuAlaLysGluGluSerValIleLeuCys	416
Db	1534	ATAGACCATCAGCATTTGTTTGGTGCAGCTCGCAAAAGAAATCAGTAATCTTATGC	1593
Qy	417	ProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSer	436
Db	1594	CAAGGAGTGTCTTGGAAATGGCAACTGGTCCGCAATTACTTTGCTTGTTCCTCT	1653
Qy	437	SerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsnLysLysAsn	456
Db	1654	TCTCTTCAAGATGCTCGGAAGGATCAAAATCATTCTGTCAAGGAAACAAGAGAGAAAT	1713
Qy	457	SerIleAsnGlyCys	461
Db	1714	TCGAGCGATGATTC	1728
RESULT 3			
LOCUS	BD012010	10966 bp	DNA linear
DEFINITION	Creation of iron-deficiency resisting rice plant.		
ACCESSION	BD012010		
VERSION	BD012010.1	GI:22092199	
KEYWORDS	WO 0101762-A/1.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 10966)		
AUTHORS	Mori,S., Nakanishi,H., Takahashi,M. and Nishizawa,N.		
TITLE	Creation of iron-deficiency resisting rice plant		
JOURNAL	Patent: WO 0101762-A 1 11-JAN-2001;		
	JAPAN SCIENCE AND TECHNOLOGY CORP, SATOSHI MORI, HIROMI NAKANISHI,		
	MICHIKO TAKAHASHI, NAKO NISHIZAWA		
COMMENT	OS Horudenum vulgare L. var. Igri		
	PN WO 0101762-A/1		
	PD 11-JAN-2001		
	PF 04-JUL-2000 WO 2000JP004425		
	PR 05-JUL-1999 JP 99P 190318		
	PI SATOSHI MORI, HIROMI NAKANISHI, MICHIKO TAKAHASHI, PI NAKO		
	NISHIZAWA		
	PC A01H5/00, C12N5/14, C12N15/52		
	CC		
FEATURES	Key	Location/Qualifiers.	
source		Location/Qualifiers	
		1..10966	
		/organism="unidentified"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:32644"	
BASE COUNT	3082 a	2422 c	2366 g 3096 t
ORIGIN			
Alignment Scores:			
Pred. No.:	2.1e-132	Length:	10966
Score:	1988..00	Matches:	457



Db 7476 CAGTCCGAGAGTGGCAAGAGCTCGGAATATTGGTATCGCTGACGAGGTTTACGGC 7535  
 Qy 261 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro 280  
 Db 7536 AACTGGTCTGGGCAAGCCCGCTTTATCCGATGGGGCTCTTTGGSCACATTTGCCCG 7595  
 Qy 281 ValLeuSerIleGlySerLeuSerIlyssSerIleValProGlyTyrArgLeuGlyTyr 300  
 Db 7596 GTCTTGTCCATTGGATCTCTGCTCAAGTGTGGATAGTCCCTGGATGGCGACTTTGGATGG 7655  
 Qy 301 ValAlaValTyrAspProThrIlystleIleuGluIlyThrLysIle- 315  
 Db 7656 GTGGCGGTGTACGCCCCCAAGATTTTGTAGAAACTTAAGT-AGCTTTAGCTCCCTA 7714  
 Qy 315 - 315  
 Db 7715 TCATTCTTCTCATATGCTACTGTGGGATTAGTATTTTGTCTAAATTTGTACTGCTTTG 7774  
 Qy 316 - - - - -SerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThr 331  
 Db 7775 TTTATTTCAGATCTCTACGCTATTACGAATTACCTTAATGTCTCAACGACCCAGCAACC 7834  
 Qy 332 PheValGln- 334  
 Db 7835 TTCGTTTCAGGTTAGTCTTTGGTCTTTCGCCCTATTTTGTCTCATGCTCCCTGTGTTCATGTC 7894  
 Qy 334 - 334  
 Db 7895 AATGACCGGCTCAAGTTAGTATATAGAGTTTGTGTAAGTGTGAATGTGCAAGTCCAA 7954  
 Qy 334 - 334  
 Db 7955 CATGATGGAAGAAAGATACATCTATTTTGTAGTCATTCGCCCTTTGTTGTTGANTCCATA 8014  
 Qy 334 - 334  
 Db 8015 AATATAATAACACAAAGCCAGAACCAACTATTTGTAATAGACTATTTTCTTAGAAATA 8074  
 Qy 334 - 334  
 Db 8075 TACATTGTATTTGAGCATGCCATATTTCTTTTCGATCAAGTATGCAATATTTAAACTT 8134  
 Qy 334 - 334  
 Db 8135 GCATTGTACTACGATATACCATGTTGTTAAGAAATTTCTTTACCTACACACCTTGTCTC 8194  
 Qy 334 - 334  
 Db 8195 GCATCTTCATATTTTGATATCCCTTGACATTAATTTGTTCTCTTATGATTACACAACTTAAT 8254  
 Qy 335 - - - - -GluAlaLeuProIlystleuGluAsnThr 344  
 Db 8255 TATGGATTTTGTGTATCAAAATGTTTAGGAGCTCTCTCTAAATTTCTTGAGAACACA 8314  
 Qy 345 LysAlaAspPhePheIlystleuLeuLeuGlySerIleGlyLeuLeuGlySerIleGlyCysTyr 364  
 Db 8315 AAGCAGATTTCTTTAAGAGGATTAATGTCTACTAAAGGAATCATCAGAGATATGTTAT 8374  
 Qy 365 ArgGluIleGlyGluAsnIlystleuThrCysProHisIlyProGluIlySerMetPhe 384  
 Db 8375 AGGGAATAAAGGAAACAAATATATTACGTGTCTTCAACGCCAGAGGATCGATGTTT 8434  
 Qy 385 ValMet- 386  
 Db 8435 GTAATGTAAGCTAAGCATAGACTTACTTTTAAAGGTTAATCTGGGATCTCAGTGCATCC 8494  
 Qy 386 - 386  
 Db 8495 AACAAACATCAAAATCAAAATATAATATGTTTGTCTATGGATCTTTTGAAGATGCATG 8554  
 Qy 386 - 386  
 Db 8555 CATTTGAAGAAATAATGAAGAGAGTTGAAATTAATTTTAGGACTAAATCTTCTTCATATCAT 8614

Qy 386 - 386  
 Db 8615 TGTCCATTTTTTTGTTATTACTGTAAATTTGGTAACTCAATATATATACAAAAGTTT 8674  
 Qy 386 - 386  
 Db 8675 CTTCCCATTTTGTAGTAAGATGACTTCCTTCTATAACCATGTATTAATCTCCATGTAAA 8734  
 Qy 387 - - - - -ValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCys 405  
 Db 8735 CAGGTCAAACTAACTTACATCTTTTGGAGGAGATCCATGACGACATAAAATTTTGTCTGC 8794  
 Qy 406 LysLeuAlaLysGluIleSerValIleLeuCysPro- 417  
 Db 8795 AAGCTGCAAGAAAGAAATCTGTAATTTTATGTCAGGTAGGAATGTATATGTCGCAATTTT 8854  
 Qy 417 - 417  
 Db 8855 AAAGGAAACTATATGGAATATAATAATATCTTCTTGTATATACTAAACAATACTTCTCCAT 8914  
 Qy 417 - 417  
 Db 8915 CTTAAATAAATGTCTTACACTTAGACAAATTTTATATAGTACTAGTACAAAGTTGAAA 8974  
 Qy 417 - 417  
 Db 8975 CAGTTATTTGGGACAGAGGAGTAGTATATATTGTTGTGAGAAACATAAGGTTATGTTGA 9034  
 Qy 417 - 417  
 Db 9035 CTGATATATGCTTCTTAAATGTGAACATGTTCTCTTAATGTTTGTGATATACGAAG 9094  
 Qy 417 - 417  
 Db 9095 TTCTTATCAGTTTCCGAGATGACTACACATAAATGATTACCATATCATTTGTCAGAAATG 9154  
 Qy 417 - 417  
 Db 9155 TATTACCACATAGAAATATCTTCTTTTATGCAAGACTAGCATGGCATGACTTTTC 9214  
 Qy 417 - 417  
 Db 9215 CTGTACTATGTGCTTTTTTTTTTCTCGTTACATGTTTGTCTTCTCACAAAATAATA 9274  
 Qy 417 - 417  
 Db 9275 ATACCAAGCACATGTTTCCAAATGATTATTAATTTTGAAGTGTTTTCAACCAACTTA 9334  
 Qy 417 - 417  
 Db 9335 TATATCTTATAGTTCTAAAAAACCCTATATATATGTTAACTTAACCTAACAAAACATTATATA 9394  
 Qy 418 - - - - -GlySerValLeuGlyMetGluAsnTyrValArgIleThrPhe 431  
 Db 9395 TGTTTTCTCTCTAATACAGGAGTGTCTTGAATGGAATAATTTGGTCCGTATTACTTTT 9454  
 Qy 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValIlysserPheCysGlnArg 451  
 Db 9455 GCTGTGTTCCATCTTCTTCAAGATGGACTCGAAAGGGTCAATCATCTTCTGTCAAGG 9514  
 Qy 452 AsnLysLysLysSerIleAsnGlyCys 461  
 Db 9515 AACAGAAGAAGAAATCTATAATGTTGT 9544

## RESULT 4

LOCUS E49893 10966 bp DNA linear  
 DEFINITION Creation of iron-deficiency resisting rice plant.  
 ACCESSION E49893  
 VERSION E49893.1 GI:18629372  
 KEYWORDS JP 2001017012-A/1.  
 SOURCE unidentified



Db 8135 GCATTGTACTACGAGTATACCATGTTGTTAAGAAATTTCTTTACCTACACACCTTGTCTC 8194  
Qy 334 ----- 334  
Db 8195 GCATCTTCATATTTTGCATATCCTTGACATTAATGTTCTCTTATGATTACACACAATTAAT 8254  
Qy 335 ----- 335  
Db 8255 TATGATTTTGTGCTATCAAAATGTTTAGGAGCTCTTCTTAAATTTCTTGAGAACACA 8314  
Qy 345 LysAlaAspPhePheLysArgLeileGlyLeuLeuLysGluSerSerGluileCysTyr 364  
Db 8315 AAAGCAGATTTCTTTAAAGAGGATTAATGCTCTACTAAAGGAATCATCAGATATGTTAT 8374  
Qy 365 ATGGluileLysGluAsnLysTyrileThrCysProHisLysProGluLysSerMetPhe 384  
Db 8375 AGGGAATAAAGGAACAATAATATTAATACGTGCTCTCAAGCCAGAGGATCGATGTTT 8434  
Qy 385 ValMet----- 386  
Db 8435 GTAATGGTAAGCTAAGCATAGACTTACTTTTAAAGTTAACTCGGATCTCAGTGCATCC 8494  
Qy 386 ----- 386  
Db 8495 AACAAACAATCAATCAAAATATAATTAATGTTTCTATGGATCTTTTGAAGATGCATG 8554  
Qy 386 ----- 386  
Db 8555 CATTTGAGAATAATGAAGAGAGTTGAATATTTTAGGACTAATCTCTCGATATCATTT 8614  
Qy 386 ----- 386  
Db 8615 TGTCATTTTTTTGTTATTACTGTAATGTAATGTAACACTCAAAATCATATTAACAAAAAGTTT 8674  
Qy 386 ----- 386  
Db 8675 CCTCCATTTTGTAGTAAGTACTCTCTTCTATACCATGTAATTAATCTCCATGAAA 8734  
Qy 387 ---ValLysLeuAsnLeuHisLeuLeuGluileHisAspAspileAspPheCysCys 405  
Db 8735 CAGGTCAAACTAAACTTACATCTTTTGGAGAGATCCATGACGACATAAATTTTGTCTGC 8794  
Qy 406 LysLeuAlaLysGluLysSerValleLeuLysPro----- 417  
Db 8795 AAGCTCGCAAGGAAGAACTGTAAATTTTATGTCAGTAGGAATGTATATGGCCATTTT 8854  
Qy 417 ----- 417  
Db 8855 AAGAGAAACTATATGGATTAATATATCTTCTTGTATATACTAAACAATCTCTCCAT 8914  
Qy 417 ----- 417  
Db 8915 CCTAAATAAATGCTTACACTAGCACAAATTTTATATAGTACTAGTACAAAGTTGAAA 8974  
Qy 417 ----- 417  
Db 8975 CAGTTATTTGGGACAGAGGAGTAGTATATATTTGTGTGAGACATAAGTTATGTTGA 9034  
Qy 417 ----- 417  
Db 9035 CTGATATATGCTTCTTAAATGTGAACATGTTCTCTTATGTTTGTGATTGTATACGAAG 9094  
Qy 417 ----- 417  
Db 9095 TTCTTATCAGTTTCCGAGATGACTACACATAAATGATTAACCATATATCTGCAGAAAATG 9154  
Qy 417 ----- 417  
Db 9155 TATTACCACATTAAGATATCTTTCTTTTATGCAAGACTAGCATGGCATGTACTTTTC 9214  
Qy 417 ----- 417  
Db 9215 CTGTACCTATGCTGCTCTTTTCTCTGTTTCTCGTTACATGTTTGTGCTTCTCACAATAATA 9274

Qy 417 ----- 417  
Db 9275 ATACCAAGCACATGTTTCCAAATGATTATTATAATTTTGGAGTGTGTTTTCACCAACTTA 9334  
Qy 417 ----- 417  
Db 9335 TATATCTTCAAGTCTTAAAAAAACCGTATATATGTTTAATCTTAACAAAAAATTTATATA 9394  
Qy 418 -----GlySerValLeuGlyMetGluAsnTrpValArgileThrPhe 431  
Db 9395 TGTCTTCTCTCTAATACAGGAGCTGTTCTTGGATGGAATTTGGTCCGTATTTT 9454  
Qy 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArg 451  
Db 9455 GCCTGGTCTTCATCTTCTCTCAAGATGACCTGAAAGGCTCAATCATCTCTCAAGG 9514  
Qy 452 AsnLysLysLysAsnSerileAsnGlyCys 461  
Db 9515 AACAGAAGAGAAATCTATAATGTTGT 9544  
RESULT 5  
LOCUS AB024006 10966 bp DNA linear PLN 27-NOV-1999  
DEFINITION Hordeum vulgare naat-B and naat-A genes for nicotianamine  
aminotransferase, complete cds.  
ACCESSION AB024006  
VERSION AB024006.1 GI:6459088  
KEYWORDS nicotianamine aminotransferase.  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poidea; Triticeae; Hordeum.  
1 (sites)  
REFERENCE  
AUTHORS Takahashi, M., Yamaguchi, H., Nakanishi, H., Shioiri, T.,  
Nishizawa, N. K. and Mori, S.  
TITLE Cloning two genes for nicotianamine aminotransferase, a critical  
enzyme in iron acquisition (Strategy II) in graminaceous plants  
JOURNAL Plant Physiol. 121 (3), 947-956 (1999)  
MEDLINE 20027437  
PubMed 10557244  
REFERENCE 2 (bases 1 to 10966)  
AUTHORS Mori, S., Takahashi, M. and Nakanishi, H.  
TITLE Direct Submission  
JOURNAL Submitted (22-FEB-1999) Satoshi Mori, University of Tokyo, Dept. of  
Applied Biological Chemistry; Yayoi 1-1, Bunkyo-ku, Tokyo 113-8657,  
Japan (E-mail:asmori@hongo.ecc.u-tokyo.ac.jp,  
Tel:81-3-3812-2111 (ex.5106), Fax:81-3-3812-0544)  
Location/Qualifiers  
1. .10966  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="genomic DNA"  
/sub\_species="Igr1"  
/db\_xref="taxon:112509"  
/clone\_lib="lambda FIXII"  
Join(650..1299,1483..1825,1922..2140,2244..2303,  
2761..2916,3263..3356,3735..3868)  
/gene="naat-B"  
Join(650..1299,1483..1825,1922..2140,2244..2303,  
2761..2916,3263..3356,3735..3868)  
/gene="naat-B"  
/codon\_start=1  
/product="nicotianamine aminotransferase"  
/protein\_id="BAA87054.1"  
/db\_xref="GI:6469089"  
/translation="MATYRQSDGVAANGLAVAAANGKSNHGVAANGKSNHGCHVD  
ADANGKSNHGVAANGKSNHGVAANGKSNHGVAANGKSNHGVAANGKSNHGCHVD  
SNEHAEKSNHGVAANGKSNHGVAANGKSNHGVAANGKSNHGVAANGKSNHGCHVD  
SVQEKGPRLVPLAHGDPSPVFAETAVEAEVAALATGQFNCYCPAGVGLPAARSA  
VAEHLQSQVYMLSDADFLTAGTQALEVLPVLAQTAGANILLPRGYPNPARAA  
FNRLVRHFDLIPDKGIDIDLSIESIADKNTAMVIINPNPCSVSVYSHLSKVAE  
VAKRLGILVIADEVYKLVLSGAPFIPMGVFGHITPVLISGLSKSVLPVGRWLVGA



QY 334 ----- 334  
 Db 8195 GCATCTTCATATTTTGATATCTTGACATATTGTTCTCTTATGATTCACACAACTTAAT 8254  
 QY 335 -----  
 Db 8255 TATGATTTTGTGTATCAAAATGTTAGGAAGCTCTCTCTAAATTTCTTGAGACACA 8314  
 QY 345 LysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 364  
 Db 8315 AAAGCAGATTCTTTAAGAGGATATTGCTCTACTAAGAGATCATCAGATATGTTAT 8374  
 QY 365 ArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPhe 384  
 Db 8375 AGGGAATAAAGGAACAAATATATTACGTGCTCTCAAGCCAGAAGGATCGATGTT 8434  
 QY 385 ValMet----- 386  
 Db 8435 GTAATGGTAAGCTAAGCATAGACTTACTTTTAAGGTTAATCTGGGATCTCAGTGCATCC 8494  
 QY 386 ----- 386  
 Db 8495 AACAAACAATCAAAATCAAAATATAATTATGTTTGTCTATGGATCTTTTGAAGATGCATG 8554  
 QY 386 ----- 386  
 Db 8555 CATTTGAAGAATAATGAAGAGAGTTGAAATATTATTTAGGACTAATCTTCTGTGATCATTT 8614  
 QY 386 ----- 386  
 Db 8615 TGTCAATTTTTTGTATTACTGTAAATGTTGAACACTCAATCATATTACAAAAGTTT 8674  
 QY 386 ----- 386  
 Db 8675 CCTCCATTTTGTAGATGACTTCCTTCTATAACCATGATTAACTTCCATGTA 8734  
 QY 387 ---ValLysLeuAsnLeuHisLeuLeuGluIleHisAspIleAspPheCys 405  
 Db 8735 CAGTCAACTAACTTACATCTTTGGAGAGATCCATGACGACATAATTTTGTCTGC 8794  
 QY 406 LysLeuAlaLysGluSerValIleLeuCysPro----- 417  
 Db 8795 AAGCTCGAAAGGAATCTGTAATTTATGTCAGGTAGGAATGTATATGGCCATTT 8854  
 QY 417 ----- 417  
 Db 8855 AAAGAAACTATATGGAAATAATAATATCTTCTTTGTTATATAACAATACTTCTCCAT 8914  
 QY 417 ----- 417  
 Db 8915 CCTAAATPAAATGCTTACACTTAGCACATTTTATAGTATAGTACAAAGTTGAAA 8974  
 QY 417 ----- 417  
 Db 8975 CAGTTATTTGGACAGAGGAGTAGTATATATTGTTGTGAGACATAAGTTATGTTGA 9034  
 QY 417 ----- 417  
 Db 9035 CTGATATATGCTTCTTAAATGTGAACATGTTCTTATGTTTGTGATGTATACGAAG 9094  
 QY 417 ----- 417  
 Db 9095 TTCTTATCAGTTTCCGAGATGACTACACATAAATGATTACCATATCATTTGTCAGAAAATG 9154  
 QY 417 ----- 417  
 Db 9155 TATTACCACATTAGATATCTTTCTTTTATGCAAGACTAGCANGGCAATGACTTTTC 9214  
 QY 417 ----- 417  
 Db 9215 CTTGTACTATGTCTTTTTTTTCTCGTTACATGTTTGTGCTTCTTCACAAAATAATA 9274

QY 417 ----- 417  
 Db 9275 ATACCAAGCACATGTTCCAAATGATTATTAAATAATTTGAGGTGTTTTTCAACCAACTTA 9334  
 QY 417 ----- 417  
 Db 9335 TATACTTTTCATAGTCTTAAAAAAACCGTATATATGTTAACTCTAAACAAAAACTTATATA 9394  
 QY 418 -----  
 Db 9395 TGTTTCTCTCTAATACAGGAGTGTCTTGGAAATGGAATAATTTGGTCCGTATATCTTT 9454  
 QY 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArg 451  
 Db 9455 GCCTGGTTCATCTTCTTCAAGATGACATCGAAAGGTCAAATCATCTTGTCTCAAGG 9514  
 QY 452 AsnLysIleLysAsnSerIleAsnGlyCys 461  
 Db 9515 AACAGAAGAAGAAATCTATAATGTTGT 9544  
 RESULT 6  
 AX653727  
 LOCUS AX653727 1290 bp DNA linear PAT 22-MAR-2003  
 DEFINITION Sequence 3597 from Patent WO03000898.  
 ACCESSION AX653727  
 VERSION AX653727.1 GI:29156541  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE  
 1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
 Katagiri, F., Quan, S., Tao, Y., Whitham, S.S., Xie, Z., Zhu, T. and Zou, G.  
 Plant genes involved in defense against pathogens  
 Patent: WO 03000898-A 3597 03-JAN-2003;  
 Syngenta Participations AG (CH)  
 FEATURES  
 Location/Qualifiers  
 1..1290  
 /organism="Oryza sativa"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4530"  
 BASE COUNT 280 a 307 c 400 g 297 t 6 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,79e-79 Length: 1290  
 Score: 1228.50 Matches: 229  
 Percent Similarity: 71.26% Conservative: 81  
 Best Local Similarity: 52.64% Mismatches: 100  
 Query Match: 51.40% Indels: 25  
 DB: 6 Gaps: 5  
 US-10-019-783-1 (1-461) x AX653727 (1-1290)  
 QY 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaValGluTrpAsnPheAlaArgGly 47  
 Db 7 GACGGTTCGGWAAMGGCG 63  
 QY 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67  
 Db 64 AACGGCGGCTTCTCGGGTGTACGGCGTG-----TCGGTGGCGGCKTGTGGCGGCGG 117  
 QY 68 IleSerAlaSerValGlu-----SerGlyProArgPro 79  
 Db 118 GTCAAGCGCGGATCG 177  
 QY 80 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 99  
 Db 178 GTGATCCCGATGGGACACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 237  
 QY 100 AlaGluAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGly 119

238 GCCGTGACCGCGTGGGGGGGCTCCGCTCGGGGAGCAGCAACTCTACTCTCTCTGCT 297  
 QY  
 120 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 139  
 Db  
 298 GTCGGCTTCGAGCTGACGAGAGTCTATCGCGGTACTTATCGGAGACTTGCATAT 357  
 QY  
 140 LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159  
 Db  
 358 GAGCTATCAGCTGATGATGATCTCACTGACAAAGTGTCTGCTCAAGCGATTGAGATCATC 417  
 QY  
 160 IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 179  
 Db  
 418 TGCTCTGCTTCTAGTCTGCG---CCTGGTGCACAACTCTGTGCCAAGCCAGGTTACTTG 474  
 QY  
 180 AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValAlaArgHisPheAspLeuIlePro 199  
 Db  
 475 TTCCACGAGGCACGCGCAGTCTTCAATGGCATGGAGTCAAGTACTTTGATCTTCTCCCA 534  
 QY  
 200 AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219  
 Db  
 535 GAGAGTGGCTGGGAGGTGATCTTCTATGGAGTGCAGGAATCTGCTGACAAAGAACACGGTT 594  
 QY  
 220 AlaMetValIleIleAsnProAsnProCysGlySerValTyrSerTyrAspHisLeu 239  
 Db  
 595 GCATGTGCTATATCAATCCAGGAAATCCCTGTGGCAATGCTACACTTCTGAGCATTTG 654  
 QY  
 240 AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyr 259  
 Db  
 655 GCCAAG-----GCCAATGAAGTGTAT 675  
 QY  
 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279  
 Db  
 676 GCACATTTGACGTTTGGCAGCAATAATTTGTGCCGATGGGGTGTGTGCTCAGTAGCT 735  
 QY  
 280 ProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrTrpArgLeuGly 299  
 Db  
 736 CCAGTCTTACGTTAGGATCATATCAAGAGATGGGTGTGCTGCTGATGGCAGCTTGA 795  
 QY  
 300 TrpValAlaValTyrAspProThrLysIleLeuGluLysThrLysIleSerThrSerIle 319  
 Db  
 796 TGGATGTGACGAGCATCTTAATGGCGTATTTCAAGAGCAAGGATGTTGGAAGCATC 855  
 QY  
 320 ThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLys 339  
 Db  
 856 CAGAGTTACCTTGATATCTCGCTGATCTTGCAACATTTATCCAGGGAGCAATCCACAA 915  
 QY  
 340 IleLeuGluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSer 359  
 Db  
 916 CTCATAGAGAACACAAAGGAAGAAATCTTCGAAAGACCCGTGATGCTTAAGCGAGACC 975  
 QY  
 360 SerGluIleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysPro 379  
 Db  
 976 GCAGATTTTGTGGGAGACTGAAGGCATCAGTTGCATCATCTTCCCAAGCAACCT 1035  
 QY  
 380 GluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGluIleHisAsp 399  
 Db  
 1036 GAGGTTCCATGTTTGTATGTTGAAACTGGATCTGCTCTGCTGCAAGGCATCAAGAT 1095  
 QY  
 400 AspIleAspPheCysLysLeuAlaLysGluSerValIleLeuCysProGlySer 419  
 Db  
 1096 GACATGGACTTCTGCTGCCACTAGCAAGAGAAATGTTGATCTTTTCCAGGATGT 1155  
 QY  
 420 ValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGln 439  
 Db  
 1156 GCTGTTGATACAAAGATTGCTCGAATCACCITCGCCATCGAAGCGTCTTCTCTTGA 1215  
 QY  
 440 AspGlyLeuGluArgValLysSerPheCysGlnArgAsnLysLys 454  
 Db  
 1216 GATGTTATTGATAGCTCAAGTCTCTGCTCGGACATAGCAAG 1260

RESULT 7  
 AX653224

LOCUS AX653224 920 bp DNA linear PAT 22-MAR-2003  
 DEFINITION Sequence 3094 from Patent WO03000898.  
 ACCESSION AX653224  
 VERSION AX653224.1 GI:29156038  
 KEYWORDS Oryza sativa  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.  
 1  
 REFERENCE  
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
 TITLE Plant genes involved in defense against pathogens  
 JOURNAL Patent: WO 03000898-A 3094 03-JAN-2003;  
 Syngenta Participations AG (CH)  
 FEATURES  
 Location/Qualifiers  
 1..920  
 /organism="Oryza sativa"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4530"  
 BASE COUNT 275 a 184 c 205 g 256 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,56e-74 Length: 920  
 Score: 1150.50 Matches: 224  
 Percent Similarity: 75.9% Conservative: 29  
 Best Local Similarity: 67.2% Mismatches: 29  
 Query Match: 48.14% Indels: 51  
 DB: Gaps: 2

US-10-019-783-1 (1-461) x AX653224 (1-920)

QY 127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspVal 145  
 Db 67 AGTGTGTGGCAGATCATTTGTACAGGCATCTCCATCAAGCTATCTTCATGACATC 126  
 QY 147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGluThr 166  
 Db 127 TTCCTAACCGCTGGAGGAACCTCAGSCCATCGAGTCTGTAATCTCAATCTTGCCTCAA 183  
 QY 167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAla 185  
 Db 184 CTTGCGACAAACATATTCCTTAGACCGCTACCCAACTATGAAGCTCGAGCGCG 243  
 QY 187 PheAsnLysLeuGluValAlaArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp 206  
 Db 244 TTCACAAACCTTGAA----- 258  
 QY 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 226  
 Db 258 ----- 258  
 QY 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246  
 Db 259 -----GTGCGAGAGGTAGCA 273  
 QY 247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 266  
 Db 274 AGGAAGCTTGGCATATTGTAATTACTCATGAGGTGTATGTAATTTGTTTGGGAGT 333  
 QY 267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer 286  
 Db 334 TCCCATTTGTCCCAATGGTGTCTTTGGGCATCTCGTACCAATATTACCATAGATCG 393  
 QY 287 LeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspPro 306  
 Db 394 CTATCAAGAGGTGGATAGTGCAGGATGGGACTTGTGTTGGTAGCAATATGTGACCCC 453  
 QY 307 ThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 326  
 Db 454 AAGAAGACTCTTACAAGAACCAAGATTGCAATTAATTACTTAATTTCTTAATTTTCA 513

```

QY 327 ThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrIysAla 346
Db 514 ACTGATCCGACACITTCATTAGGAGCTCTACCGAATATCTTAAGAATACCAAGGAA 573
QY 347 AspPhePheLysArgIleLeuGlyLeuLeuLysGluSerSerGluIleCysTyrArgGlu 366
Db 574 GAATTCCTTTAAGAGGATAATTCATTCAGGAGCTCTACCGAATATCTTAAGAATACCAAGGAA 573
QY 367 IleLysGluAsnLysTyrIleThrCysProHisLysProGluLysSerMetPheValMet 386
Db 634 ATAAAGGATATTAATGCATCACTTGTCTCTCAAGCCGGAAGGATCATGTTGTGTGATG 693
QY 387 ValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLys 406
Db 694 GTCAAAATTCACCTATATCTTTGGAGGAATCCATGATGATGTTGATTGTTGCCAA 753
QY 407 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrp 426
Db 754 CTTCGGAAGAAGAGTCGGTGATCTTTGCCAGGAGTGCTCGGAATGAAGAATTGG 813
QY 427 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLys 446
Db 814 GTTCGCAATACCTTTGCTATTGATTCATCTCTCTCTGGATGGTCTTGAGAGGATCAA 873
QY 447 SerPheCysGlnArgAsnLysLysLysAsnSerIleAsn 459
Db 874 TCCTTCTGCCAAGGACACAAAGAAAACCCCTTTAAAT 912

RESULT 8
AX575731
LOCUS AX575731 1389 bp DNA linear PAT 07-JAN-2003
DEFINITION Sequence 9 from Patent WO02072848.
ACCESSION AX575731
VERSION AX575731.1 GI:27552220
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1

REFERENCE
AUTHORS Badur R. and Geiger M.
TITLE Increase in the vitamin E content in organisms due to an increase
JOURNAL in the tyrosine aminotransferase activity
Patent: WO 02072848-A 9 19-SEP-2002;
Surgene GmbH & Co. KGAA (DE)
FEATURES
source
1. 1389
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
1. 1389
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD60072.1"
/db_xref="GI:27552221"
/translation="MSSEQQHANLAVPAFKTEKDPVTQTONGQSSVRFGGSKAAKA
STVLRGVIYMLFNCVSKVNKILPLHGDPSPVPCFRTCEAEADAVDVLRSKGN
SYGAGILPARRAAYSGVEVKFDLLPEKEWIEDLEGIEALDEKTVAMVWVNNPCG
NVASHDLKKVAETARKLGMVTSDEVDRTIPGDNFVPMGPFASVPLVTLGASK
FWVPGWKI GWIALNDPEGVFETTKVLSQIKONLDVTPDPAITIIQALPALLEKADKN
FFAKNKILKHNVDLVCDRLKDI PCVVCPPKPSCYLLTLKLSLMDNKKDIDFCV
KLAEENLVLPGDALGLKRWTRITITVEAHMLEDALERLKGFCTRHAKKTETETSSL
QALKLSNNLEM"
BASE COUNT 351 a 322 c 336 g 340 t
ORIGIN

Alignment Scores:
Pred. No.: 6.75e-72 Length: 1389
Score: 1128.00 Matches: 211
Percent Similarity: 66.52% Conservative: 87

```

```

Best Local Similarity: 47.10% Mismatches: 138
Query Match: 47.20% Indels: 12
DB: 6 Gaps: 4
US-10-019-783-1 (1-461) x AX575731 (1-1389)
QY 8 HisGlyGluAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAla 27
Db 19 CACCCCAATCTAGCGGTCCCGGTTTAAAACTGAGAAAGATCCCGTAAGCAACCAA 78
QY 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaValGluTrpAsnPheAlaArgGly 47
Db 79 AATGGTCAAAAGTAGC-----CTTTGGCGTTTCGGT---GGA 111
QY 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67
Db 112 AGTGAT-----AAGCAGCGCAAGCATCCACCGTAACGCTAGAGGTGTCACTCATG 165
QY 68 IleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
Db 166 CTCTTCGACAACTGCACCAAGAGCTCAATAAGACCATTTTACCCCTCGGCCACGGTGAC 225
QY 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107
Db 226 CCTTCGCTCTACCGTTCCTGCTACCTGTATCGAAGCTGAAGACGCCGCTCGCGAGTC 285
QY 108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
Db 286 CTTCGCTCCGCAAGGCAATCTTACGCTCCGAGCTGGGATTCCTCCCGAAGACGA 345
QY 128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
Db 346 GCGGTGCTGATATATGAACCGAGATCTCCGACAGTTAACGCCCGCAAGATATTTT 405
QY 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAla 167
Db 406 CTACCGCTGGATGCAACCAAGGATAGATGCTGTTCGAATCGTTGGTCGA---CCA 462
QY 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaPhe 187
Db 463 AACGCAAAACATCTTGTCTCCACGCTCTGCTCCCTCATTCAGCCTCGTGTCTTAC 522
QY 188 AsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIle 207
Db 523 AGTGGTCTGAGGTTCGCAAGTTTGTATCTTCTCCGAGAAAGATGGGAGATTGATCTT 582
QY 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 227
Db 583 GAAGGTATCGAAGCCATTGCAGACGAGAAACTGTGGCTATGGTGTAAATTAACCCCAAC 642
QY 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247
Db 643 AATCCCTGTGGAATGTCTACTCTCAGGACCATCTCAAAAAGGTTGCAGACGCGTAGG 702
QY 248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267
Db 703 AGCTCGGATATGTTGATCTCAGACGAAGTATATATGACCGAATATATTCGAGACAAAT 762
QY 268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
Db 763 CCATTGTTCCATGGGGAAGTTGCTTCGATAGTCCCTGTATTACACTAGACGCAATA 822
QY 288 SerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThr 307
Db 823 TCTAAGGGATGGGTTGTTCTCTGATGAAATGCGTGGATTCCTGATCATCCATCATCCGAG 882
QY 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
Db 883 GCGGTTTCGAGACCCACCAAGGTGTACATCCATCAACAGAAATCTTTCAGCTAACTCCT 942
QY 328 AspProAlaThrPheArgGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347
Db 943 GACCCCTGCCAATAATTCAGGCTGCATCTCCAGCGATCTTCGAGGATCTTCGAGAAAGCGACAAAAC 1002

```





QY 148 LeuThrAlaGlyCylThrGlnAlaIleGluValIleProValLeuAlaGlnThrAla 167  
 Db 492 CTGACCGCTGGATGCAACCAAGGATAGAGATCGTTCGAATCGTTGGCTCGA--CCA 548  
 QY 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187  
 Db 549 AACGCAACATCTTGTCCACGCTCTGCTCCCTCACTACGACGCTCGTCTGCTTAC 608  
 QY 188 AsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIle 207  
 Db 609 AGTGGTCTGAGGTTCGCAAGTTTGAATCTTCCGGAAGAAGATGGAGATTGATCTT 668  
 QY 208 AspSerLeuGluSerIleAlaAspLysAsnThrAlaMetValIleIleAsnProAsn 227  
 Db 669 GAAGGTATCGAAGCATTCACACGAGAACACTGTGGCTATGTTGTTAATTAACCCCAAC 728  
 QY 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247  
 Db 729 AATCCTGTGGAAATGTCCTCTCACGACCATCTCAAAAGAGGTTCAGACGCGCTAGG 788  
 QY 248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267  
 Db 789 AAGTCGGGATAATGTCATCTCACAGCAAGTATATGACCGACTATATTTCGGAGACAA 848  
 QY 268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287  
 Db 849 CCATTGTTTTCATCGGAAGTTTCTTCGATAGTCCCTGTATTGACACTAGCAGGCATA 908  
 QY 288 SerLysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThr 307  
 Db 909 TCTAAGGATGGTGTTCCTGGATGGAATGCTGGATGCGCTTGAATGATCCCGAG 968  
 QY 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327  
 Db 969 GCGGTTCGAGACCAACCAAGGTGTACATCCATCAACAGAACTCTTGACGTAACTCT 1028  
 QY 328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347  
 Db 1029 GACCTCGCAACAATTAATTCAGCTGCATCTCCAGCATCCTGGAGAAAGCGGACAAAC 1098  
 QY 348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIle 367  
 Db 1089 TTCCTTGCAGAGAGAACAGACATCTCAACATAATGTTGATTTGGTGTGTGATAGGCTC 1148  
 QY 368 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 387  
 Db 1149 AAGGACATCCCTGTGCTGTCTGCCAAGAAACCTGAGTCTTGCACCTTACTATTGACA 1208  
 QY 388 LysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysLeu 407  
 Db 1209 AAGTTGGAGCTGTCATTGATGATAATATCAAGACCATATAGATTTTTCGTTAAACTG 1268  
 QY 408 AlalysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427  
 Db 1269 GCCAGAGAGGAGAAATCTGCTTTCTACACGGGATGCTCTGGGTTTGAAGACTGGGTG 1328  
 QY 428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447  
 Db 1329 AGGATACCATCGAGTCGAGCTCATATGCTTGGAGTGCATCTGTGAGACTCAAGGCT 1388  
 QY 448 PheCysGlnArgAsnLysLysLys 455  
 Db 1389 TTCGTACACGTGTCACCAAGAG 1412  
 RESULT 10  
 AF301900  
 LOCUS  
 DEFINITION Arabidopsis thaliana 1611 bp mRNA linear PLN 03-DEC-2000  
 mRNA, complete cds.  
 ACCESSION AF301900  
 VERSION AF301900.1 GI:11527942  
 KEYWORDS Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

## REFERENCE

1 (bases 1 to 1611)

## AUTHORS

Gopalraj M. and Olszewski, N.E.

## TITLE

The ROOTY/SUPERROOT1 gene of Arabidopsis encodes a putative

## JOURNAL

tyrosine aminotransferase

## REFERENCE

2 (bases 1 to 1611)

## AUTHORS

Gopalraj M. and Olszewski, N.E.

## TITLE

Submitted (02-SEP-2000) Plant Biology, University of Minnesota,

## JOURNAL

1445 Gortner Ave, St. Paul, MN 55108, USA

## FEATURES

Location/Qualifiers

## source

1..1611

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/notes="ecotype: Wassilewskija Ws-2"

## gene

1..1611

/gene="RTY/SUR1"

56..1444

/gene="RTY/SUR1"

/notes="putative tyrosine aminotransferase"

/codon\_start=1

/product="ROOTY/SUPERROOT1"

/protein\_id="AAG37063.1"

/db\_xref="GI:11527943"

/translation="MSEEQPHANLAVPAFKEDPTTQTRNGSSVVRFGSDKAAKA

STVTLRGVYMLFDNGKDVNTILPLGHSDPSVYPCFRTCEAEADAVVDVLRSGKGN

SYGPGAGILPARRAVADYNNRDLPHKLTPEIDFLTAGCNGQTEIVFESLARPNANILL

PRPGPHYDARAAYSQLEVRKFDLLPEKEWIDLEGIEALADEKTVAMVIVNPNPCG

NYVSHDLKVAETARKLGMVTSDEVDRTIFGNPFVPMKGFASIVPVLILAGLSK

GMVFPWKIHWIALNDPEGVFETTKVLSIKNDLVDPPATIIQALPAILEAKDN

FRANKNLKHNVDLVCDRLKDPVCPVKPKPESCTYLLTLELSMDMKDDIDFCV

KLAENLVFLPDGALGLKNWRITIGVAHMLEDALEALERLKGCTRHAKKTETETESL

QALKLSDNNLEM"

154

/gene="RTY/SUR1"

/notes="produces a stop codon which truncates the protein

at amino acid 32 in the afl allele"

/replaces="a"

185..235

/gene="RTY/SUR1"

/notes="deleted in the hls3 allele"

/replaces=""

BASE COUNT 466 a 352 c 374 g 419 t

ORIGIN

## Alignment Scores:

Pred. No.: 9,66e-72 Length: 1611

Score: 1127.00 Matches: 211

Percent Similarity: 66.74% Conservative: 89

Best Local Similarity: 47.10% Mismatches: 137

Query Match: 47.15% Indels: 12

DB: 8 Gaps: 4

US-10-019-783-1 (1-461) x AF301900 (1-1611)

QY 8 HisGlyGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 27

Db 74 CACGCCAATCTAGCGGTCCCGCTTTAAACTGAGAAAGATCCCATACGCAACACGA 133

QY 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 47

Db 134 AATGTCAAAGTAGC-----GTTTGGCGTTTCGT---GGA 166

QY 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67

Db 167 AGTGAT-----AAGCGCAGGAAGCATCCACCGCTTAGAGGTGTCTACATG 220

QY 68 IleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87



Db 74 CAGGCCAATCTAGCGTTCCCGGTTTAAACTGAGAAAGATCCCGTAACGCAACGCAA 133  
 Qy 28 AsnGlyLysSerAsnGlyHisAlaAlaValGluTrpAsnPheAlaArgGly 47  
 Db 134 AATGGTCAAGTAGC-----GTTGGCGTTTCGGT---GGA 166  
 Qy 48 LysAspGlyLeuLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67  
 Db 167 AGTGAT-----AAGCGCGAAAGCATCCACCGTAACGCTTAGAGGTGTCACTACATG 220  
 Qy 68 IleSerAlaSerValGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87  
 Db 221 CTCTTCGCAACTCGCGCAAGACAGCATCAATAGACCAATTTTACCCCTCGGCCACGGTGAC 280  
 Qy 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107  
 Db 281 CTTTCCTCGTCTACCCCTGCTTCGGTACCTGTATCGAAGCTGAAGACGCGCTGTCGACGTC 340  
 Qy 108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127  
 Db 341 CTTCGCTCGGCAAGGCAATCTTACGGTCCCGGAGCTGGGATCTCCCGCGAAGACGA 400  
 Qy 128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147  
 Db 401 CGCGTGTGCTGATATATACCGAGATCTTCGCACAAAGTTAACGCCCTGAAGATATTTT 460  
 Qy 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuLeuAlaGlnThrAla 167  
 Db 461 CTGACCGCTGATGTCAACACAGGATAGATCGTTCGATCGTTCGCTCGCA---CCA 517  
 Qy 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187  
 Db 518 AACGCAACATCTTCTCCACGCTCCGCTCCCTCATTCAGCAGCTCGTGTCTGTAC 577  
 Qy 188 AsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIle 207  
 Db 578 AGTGCTCGAGGTTCGAAGTTGATCTTCTCCGAGAAAGATGGAGATTGATCTT 637  
 Qy 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 227  
 Db 638 GAAGTATCGAAGCCATTGCAGACGAGAAACTGTGGCTATGGTTGTATTAATCAACCCCAAC 697  
 Qy 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247  
 Db 698 AATCCCTGTGAATGTCTACTCTCACACCATCTCAAAAAGTTGCAGACACGCTAGG 757  
 Qy 248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267  
 Db 758 AAGCTCGGATAATGGTGATCTCAGCAAGATATATGCCAATATATTCGGAGACAAT 817  
 Qy 268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287  
 Db 818 CCAATTGTTCATGGGAAGTTTGTCTCGATAGTCCCTGTATTGACACTAGCAGGCATA 877  
 Qy 288 SerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThr 307  
 Db 878 TCTAAGGATGGGTGTTCCTGGATGAATAATGGCTGGATCGCTTGAATGATCCCGAG 937  
 Qy 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327  
 Db 938 GCGGTTTTCGAGACCAACCAAGGTGTACATCCATCAACAGATCTTTCAGTAACTCTCT 997  
 Qy 328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347  
 Db 998 GACCTCGCAATAATTCAGGCTGCATCTCCAGCGATCTCGGAGAAAGCGGACAAAC 1057  
 Qy 348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIle 367  
 Db 1058 TTCTTTCGAAGAAGACAAGATCTCAACATATGTTGATTTGGTGTGTATAGGCTC 1117  
 Qy 368 LysGluAsnLysTyrIleThrCysProHisLysProGluLysMetPheValMetVal 387  
 Db 1118 AAGGACATCCCTGTGTCTGTCTGCCAAGAAACCTTGAGTCTTGCACTTACTTATTGACA 1177

Qy 388 LysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysLeu 407  
 Db 1178 AAGTTGGAGCTGTCATTTGATGATAATATCAAGACCATATAGATTTTCGCTAAGCTG 1237  
 Qy 408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427  
 Db 1238 GCCAGAGAGAGAATCTCGTGTTCCTACCAAGGTATGCTCTGGGTTTGAAGAACTGATG 1297  
 Qy 428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447  
 Db 1298 AGATATACCATCGAGTCGAGCTATATCTTGAGATGCCACTTGAGAGACTGAAGGT 1357  
 Qy 448 PheCysGlnArgAsnLysLysLys 455  
 Db 1358 TTCTGTACGCTCATGCCAAGAAG 1381  
 RESULT 12  
 AX507289  
 LOCUS AX507289 1389 bp DNA linear PAT 27-SBP-2002  
 DEFINITION Sequence 1984 from Patent WO0216655.  
 ACCESSION AX507289  
 VERSION AX507289.1 GI:23388526  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1  
 REFERENCE Harper, J.F., Kreps, J., Wang, X. and Zhu, T.  
 Stress-regulated genes of plants, transgenic plants containing  
 same, and methods of use  
 Patent: WO 0216655-A 1984 28-FEB-2002;  
 The Scripps Research Institute (US); Syngenta Participations AG  
 (CH)  
 FEATURES  
 Location/Qualifiers  
 1..1389  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3702"  
 BASE COUNT 389 a 322 c 337 g 341 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1 31e-71 Length: 1389  
 Score: 1124.00 Matches: 211  
 Percent Similarity: 66.52% Conservative: 87  
 Best Local Similarity: 47.10% Mismatches: 138  
 Query Match: 47.03% Indels: 12  
 DB: 6 Gaps: 4  
 US-10-019-783-1 (1-461) x AX507289 (1-1389)  
 Qy 8 HisGlyGluAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAla 27  
 Db 19 CACGCCAATCTAGCGTTCCCGGTTTAAACTGAGAAAGATCCCAACGCAACCA 78  
 Qy 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGluTrpAsnPheAlaArgGly 47  
 Db 79 AATGGTCAAGTAGC-----GTTGGCGTTTCGGT---GGA 111  
 Qy 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLys 67  
 Db 112 AGTGAT-----AAGCGCGAAAGCATCCACCGTAACGCTTAGAGGTGTCACTACATG 165  
 Qy 68 IleSerAlaSerValGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87  
 Db 166 CTCTTCGCAACTCGCGCAAGACAGCATCAATAGACCAATTTTACCCCTCGGCCACGGTGAC 225  
 Qy 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107  
 Db 226 CTTTCCTCGTCTACCCCTGCTTCGGTACCTGTATCGAAGCTGAAGACGCGCTGTCGACGTC 285

```

108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
Db      CTTTCCTCCGCAAGCAAGCAATCTTACGTCCTCCGAGCTGGATCTCTCCCGCAAGACGA 345
QY      128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
Db      GCGGTGCTGATTATATGAACGAGATCTCCGACCAAGTTAAGCCTGAAGATATTTT 405
QY      148 LeuThrAlaGlyGlyThrGluAlaIleGluValIleleProValLeuAlaGlnThrAla 167
Db      CTGACCGCTGATGCAACCAAGGATAGAGATCTGTTCGAATCGTTGGCTCGA---CCA 462
QY      168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187
Db      AACGCAACATCTTGCTCCACGCTCTGGCTTCCCTCACTACGACGCTGCTGCTTAC 522
QY      188 AsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIle 207
Db      AGTGTCTCGAGGTTGCAAGTTTCACTTCTCCGAGAAAGAAATGGGAGATTGATCTT 582
QY      208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 227
Db      GAAGTATCGAAGCATTGACAGACGACACTGTGCTATGTTGTAATTAACCCCAAC 642
QY      228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247
Db      AATCCCTGTGGAATGTCTACTCTCAAGACCATCTCAAAAGGTTGCAGAGACGGCTAGG 702
QY      248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267
Db      AAGTCTCGGATAATGGTGATCTCAGACGAAGTATATGACCACTATATTCGGAGACAA 762
QY      268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
Db      CCAATTGTTCAATGGGAAGTTGCTCGATAGTCCCTGATTGACACTAGCAGGCATA 822
QY      288 SerLysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThr 307
Db      TCTAAGGATGGGTTGTTCTTGGATGAAATGGCTGGATCGCTTGAATGATCCCGAG 882
QY      308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
Db      GCGGTTTCGAGACCAACCAAGTGTACATCCATCAACAGATCTTGACGTAACCTCT 942
QY      328 AspProAlaThrPheValGlnAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347
Db      GACCTGCGCAATAATTCAGGCTGCATCTCCAGCGATCTCGGAGAAAGCGGACAAAAC 1002
QY      348 PhePheLysArgIleIleGlyLeuLysGluSerSerGluIleCysTyrArgGluIle 367
Db      TTCCTTTCGAAGAGAAACAAGATCTCAACATATGTTGATTGGTGTGTATAGGCTC 1062
QY      368 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 387
Db      AAGGACATCCCTGGTGTCTGTGTCACCAAGAACTGAGTCTTGCACCTTACTATTGACA 1122
QY      388 LysLeuAsnLeuHisLeuLeuGluLysIleHisAspIleAspPheCysCysLysLeu 407
Db      AAGTTGGAGCTGTCTTATGATGATAATATCAAGACGATATAGATTTTTCGCTAAAC 1182
QY      408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427
Db      GCCAGAGAGGAATCTCGTGTTCACAGGGATGCTCTGGGTTTGAAGAACTGGATG 1242
QY      428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447
Db      AGGATAACCATCGGAGTCCAAAGCTCATATGTTGAGGATGCATTCAGAGACTGAAGG 1302
QY      448 PheCysGlnArgAsnLysLysLys 455
Db      TTCGTACAGCTCATGCCAAGAG 1326

```

## RESULT 13

AY091293

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gene

CDS

AY091293 1420 bp mRNA linear PLN 18-SEP-2002

Arabidopsis thaliana putative tyrosine aminotransferase (At2g20610)

mRNA, complete cds.

AY091293 GI:20259032

FLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1420)

Yamada, K., Bann, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,

Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,

Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,

Carrincci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,

Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,

Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Arabidopsis Open Reading Frame (ORF) Clones

Unpublished

2 (bases 1 to 1420)

Yamada, K., Bann, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,

Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,

Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,

Carrincci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,

Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,

Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission

Submitted (20-MAR-2002) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN

Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Saik, Stanford, PGEC (SSP) Consortium members constructed and

sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K.,

Bann, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,

Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,

Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H.,

Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J.,

Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,

Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally

to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP

/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis

genome submitted to GenBank.

Location/Qualifiers

1..1420

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="2"

/clone="U10872"

/note="This clone is in pUNI 51."

ecotype="Columbia"

1..1420

/gene="At2g20610"

1..1389

/gene="At2g20610"

/codon\_start=1

/evidence=experimental

/product="putative tyrosine aminotransferase"

```

/protein_id="AAW14232.1"
/db_xref="GI:20259033"
/translation="MSEGGPHANLAVPAFKTEKEPITOTRNGQSSVWFSGDGAAKA
SVTLRGVIMLFNDCGQVNMKTLPLGHGDPVSVPFCRTCEIARDAVDVLRSGKN
SYGPGAGILPARAVADYVNRDLPHKLEFELTAGCNQGLEIUFESLAPNANILL
PRGPHVDARAAVSGLEVRKEDLLPEKEWEIDLEGIAIADENIVANVNPNNPG
NYSYDHLKVAETARKGIMWISDEVDRITFGDNPFVSMGKFAFVIVLTLAGISK
GWVPGWKIHALNDPGVFETTKVLSIKONLDVTPDPATIIQAALPAILEADKN
FKAKNKLKHNVLDVCELDKIPCVGPKPESCTYLLKLESLMDNIKDDIDFCV
KLARENILVFLPGDALGLKNWRMITIGVEAHMLEDALERLKGFCFTRHAKTETESL
QALKLSDNNLEN"
1390..1420
/gene="At2g29610"
3'UTR 402 a 326 c 346 g 346 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 1,34e-71 Length: 1420
Score: 1124.00 Matches: 211
Percent Similarity: 66.52% Conservative: 87
Best Local Similarity: 47.10% Mismatches: 138
Query Match: 47.03% Indels: 12
DB: 8 Gaps: 4
US-10-019-783-1 (1-461) x AY091293 (1-1420)
QY 8 HisGlyGluAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAla 27
DB 19 CACGCCAATCTAGCGGTGCCCGGTTTAAACTGAGAAGAGCCCATTAACCAACACGA 78
QY 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGluTyrAsnPhAlaArgGly 47
DB 79 AATGTCACAAAGTAGC-----GTTGCGGTTTCGGT---GGA 111
QY 48 LysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrIys 67
DB 112 AGTGAT-----AAGCGACGAAACATCCACCGTACAGCTTAGAGGTGTCTATCATG 165
QY 68 IleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
DB 166 CTCCTCGACACTCGCGGAAGACGCTCAATAGACCATTTTACCCTTCGGCCACGGTGAC 225
QY 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107
DB 226 CTTCCGCTACCTTGTCTCGTACCTGTATCGAAGCTGAAGACGCGCTGTCGACGTC 285
QY 108 LeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
DB 286 CTTGCTCCGCAAGGCAATCTTACGGTCCGAGAGCTGGGATTTCTCCCGCAAGACGA 345
QY 128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
DB 346 GCGGTGCTGATTATATGACCGAGATCTTCGACACAGTTTACGCTGAAGATATTTT 405
QY 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleleProValLeuAlaGlnThrAla 167
DB 406 CTGACCGCTGGATGCAACCAAGGATAGAGATCGTGTTCGAATCGTTGGTTCGA---CCA 462
QY 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187
DB 463 AACGCAACATCTGTCTCCACGCTCTGCTTCCTCCTACACGAGCTGCTGCTGCTTAC 522
QY 188 AsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIle 207
DB 523 AGTGCTCTCGAGGTTCGCAAGTTTGATCTTCTCCGAGAAAGAAATGGGATGTGATCTT 582
QY 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaValIleIleAsnProAsn 227
DB 583 GAAGGTATCGAAGCATTCGACGAGCAACACTGTGCTATGTTGTAATTAACCCCAAC 642
QY 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247
DB 643 AATCCCTGTGAAATGTCTACTCTCACGACCATCTCAAAAGGTTTCGAGACCGGTAGG 702

```

---

```

248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267
DB 703 AAGCTCGGATATGGAATGATCTCAGACGAAGTATATGACCAACTATATTTCGAGACAAT 762
QY 268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
DB 763 CCATTGTTTCAATGGGAAGTTTCTTCGATAGTCCTCTATTGACACATGACAGGCATA 822
QY 288 SerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspProThr 307
DB 823 TCTAAGGATGGGTGTTCTCTGATGAAATTTGGCTGGATCGCTTGAATGATCCCGAG 882
QY 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
DB 883 GCGGTTTTTCGAGACCAACCAAGGTGTACCAATCCATCAACACAGAAATCTTGACGTAACCTCT 942
QY 328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347
DB 943 GACCTGCCCAATAATTCAGGCTGCACATTCACGACGATCTCGGAGAAAGCGGACAAAAC 1002
QY 348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrAspGluIle 367
DB 1003 TCTTTGCAAGAGAACAAGACTCAACATAATGTTGATTGGTGTGTGATAGCTC 1062
QY 368 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 387
DB 1063 AAGGACATCCCTGCTGCTCTGCCAAGAAACCTGAGTCTTGCACTTACTTATTGACA 1122
QY 388 LysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeu 407
DB 1123 AAGTTGAGCTGTCATTGATGGATATATCAAGACGATATAGATTTTTCGTAATAACTG 1182
QY 408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427
DB 1183 GCCAGAGAGGAGAAATCTCGGTTTTCACAGGGGATGCTCTGGTGTGAAGACTGGATG 1242
QY 428 ArgIleThrPheAlaCysValProSerSerLeuGluAspGlyLeuGluArgValLysSer 447
DB 1243 AGGATTAACCATCGGATCGAAGCTCATATGCTTGGAGTGCACCTTGAGAGACTGAAGGT 1302
QY 448 PheCysGlnArgAsnLysLysLys 455
DB 1303 TTCTGTACACGTCATGCCAAGAG 1326
RESULT 14
AF301898 1611 bp mRNA linear PLN 03-DEC-2000
LOCUS Arabidopsis thaliana Columbia ROOTY/SUPERROOT1 (RTY/SUR1) mRNA,
DEFINITION complete cds.
ACCESSION AF301898
VERSION AF301898.1 GI:11527938
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 1611)
Gopalraj,M. and Olszewski,N.E.
The ROOTY/SUPERROOT1 gene of Arabidopsis encodes a putative
tyrosine aminotransferase
Unpublished
REFERENCE 2 (bases 1 to 1611)
Gopalraj,M. and Olszewski,N.E.
Direct Submission
TITLE Submitted (02-SEP-2000) Plant Biology, University of Minnesota,
JOURNAL 1445 Gortner Ave, St. Paul, MN 55108, USA
LOCATION/Qualifiers
FEATURES
1..1611
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"

```

```

/chromosome="II"
/map="between PHVB and CHL2"
/notes="ecotype: Columbia"
1..1611
/gene="RTY/SUR1"
CDS
56..1444
/notes="putative tyrosine aminotransferase"
/codon_start=1
/product="ROOTV/SUPERROOT1"
/protein_id="AAC37061.1"
/db_xref="GI:11527939"
/translation="MSEEPHANLAVPAFKTEKEPIQTRNGOSSVWRFGSGDKAAGA
SVTLRGVYVFLMNCQKVNKTLPLGHGDPSPYPCFRTCEAEDAVVDVLSGKGN
SYTGALGLPARRAVADYMRDLPHKLTPEIDFLTAGNOCIIVIFESLARPNANILL
PRGFGPHVDARAASGLVEKFDLLPEKEWIDLEGEIADENTVAMVVPVLTAGTSK
NVYSHDLKKVAETARBLGMVTSDEVYDRTIIGDNPFVSMGKFASTVPVLTAGTSK
GWVPGKLGMIANDSEGVEFTKVLQSKQNLDTVPDPAITIQALPALLEKADEN
FFAKNKLHNDVLDVCDRLKIDPCVVCPPKPSCTYLLKLELSMDNLIKIDIFCV
KLARENLVFLPGDALGLKNWRITIGVEAHMLEDALERLJKGFCTRHAKKTETETESL
QALKLSDNNLEM"
399
/variation
/gene="RTY/SUR1"
/notes="Arg changed to Gin in sur1-4"
/replacement="a"
692
/variation
/gene="RTY/SUR1"
/notes="Pro changed to Ser in rty-1"
/replacement="t"
1253
/variation
/gene="RTY/SUR1"
/notes="Leu changed to Phe in sur1-3"
/replacement="t"
1269
/variation
/gene="RTY/SUR1"
/notes="Gly changed to Gin in sur1-1"
/replacement="a"
465 a 352 c 375 g 419 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 1.59e-71 Length: 1611
Score: 1124.00 Matches: 211
Percent Similarity: 66.52% Conservative: 87
Best Local Similarity: 47.10% Mismatches: 138
Query Match: 47.03% Indels: 12
DB: 8 Gaps: 4
US-10-019-783-1 (1-461) x AF301898 (1-1611)
QY 8 HisGlyGluAlaAlaAlaAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAla 27
Db 74 CACGCCAATCTAGCGGTCCCGCGTTTAAACTGAGAAAGAGCCATACCCAAACACGA 133
QY 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaAlaAlaValGluTrpAsnPheAlaArgGly 47
Db 134 AATGTCACAAAGTAGC-----GTTTGGCGTTTCGGT---GGA 166
QY 48 LysAspGlyIleLeuAlaThrThrClyAlaLysAsnSerIleArgAlaIleArgTyrLys 67
Db 167 AGTGAT-----AAGGACGAGAAAGCATCCCGTAACGCTTAGAGGTGCTCTCATG 220
QY 68 IleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
Db 221 CTCCTCGCAACTCGCGCAAGACGCTCAATAAGACCATTTTACCCCTCGGCGACGGTGAC 280
QY 38 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107
Db 281 CCTTCCTGCTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 340
QY 108 LeuArgThrGlyGlnPheAsnGlyTyrAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
Db 341 CTTCTCGTCCGCAAGGCAATCTTACGGTCCCGGAGCTGGGATCTCCCGGCAAGACGA 400
128 AlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147
401 GCCGTTGCTGATTATATCAACCGAGATCTTCGCGACAAAGTTAAGCCCTGAAGATATTTT 460
148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThrAla 167
461 CTCACCGCTGGATGCAACAGGGGATAGATGCTTCGAATCGTGGCTCGA---CCA 517
168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187
518 AACGCAACAATCTTCCTCCCGCTCTCGCTCCCTACTACGACGCTGCTGCTGCTTAC 577
188 AsnLysLeuGluValArgHisPheAspLeuLeuProAspLysGlyTyrGluIleAspIle 207
578 AGTGTCTCGAGTTTCGAAGTTTGTATCTTCTCCGAGAAAGAAATGGAGATTGATCTT 637
208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 227
638 GAAGTATCGAAGCCATTGCGACGAGAACACTGTGGCTATGGTTGTAATTAACCCCAAC 697
698 AATCCCTGTGGAATGTCTACTCTCAGACCATCTCAAAAAGTTGCAGACGCTAGG 757
248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267
758 AAGTCGGGATAATGTTGATCTCAGACGAAGTATATGACCAACTATATTCGGAGACAT 817
268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
818 CCATTGTTCAATGGGAAGTTTCCTTCGATAGTCCCTGTATTGACACTAGCAGGCATA 877
288 SerLysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThr 307
878 TCTAAGGATGGGTGTTCTCGATGGAATTTGGCTGGATCGCTTGAATGATCCCGAG 937
308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
938 GCGGTTTCGAGACCAACCAAGGTGTACATCAACAGAAATCTTGACGTAACCTCT 997
328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347
998 GACCTCGCCACAATAATTCAGCTGCTCTCCAGCATCTCCAGGAAAGCGGACAAAAC 1057
348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIle 367
1058 TTCTTTGCAAGAAAGAACAGATACTCAACATAATGTTGTTGGTGTGTAGGCTC 1117
368 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 387
1118 AAGGACATCCCTGTGCTGCTGTCCCAAGAAACCTGAGTCTTGACATCTACTTATGACA 1177
388 LysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCysLysLeu 407
1178 AAGTTCGAGCTGCTGATGATGATATATCAAGGACGATATAGATTTTTCGTAAACTG 1237
408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427
1238 GCCAGAGAGAGAAATCTCGTGTTCACCGGGGATGCTCTGGTTCGAAAGAACTGGATG 1297
428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447
1298 AGATAACCATCGGATCGAAGCTCATATGCTTGAGGATGCACTTGAGAGACTGAAGGT 1357
448 PheCysGlnArgAsnLysLysLys 455
1358 TTCTGTACAGCTCATGCCAAGAG 1381
RESULT 15
AY050987
LOCUS
DEFINITION
Arabidopsis thaliana putative tyrosine aminotransferase (At2g20610)
AY050987 1683 bp mRNA linear PUN 18-SEP-2002

```

```

mRNA, complete cds.
AY050987
VERSION AY050987.1 GI:15293106
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 1683)
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,
Lin,J., Miranda,M., Narusaka,M., Narusaka,E., Lam,B.,
Sato,M., Seki,M., Shinn,P., Nguyen,M., Nguyen,M., Palm,C.J., Sakurai,T.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 1683)
Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M.,
Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L.,
Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,
Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,
Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and
Theologis,A.
Direct Submission
Submitted (07-AUG-2001) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL CDNA: 'RIKEN
Arabidopsis Full-Length cDNA': Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X.,
Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D.,
Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H.,
Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E., Lam,B.,
Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
FEATURES
source
1. 1683
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAFL07-08-108 (R10872)"
/note="This clone is in a modified pBluescript vector
(FLC-1) as a BamHI/XhoI insert.
ecotype: Columbia"
gene
1. 1683
/gene="At2g20610"
5'UTR
1. 87
/gene="At2g20610"
misc_difference
2
/gene="At2g20610"
/misc_difference
21. 22
/gene="At2g20610"
/note="not present in genomic sequence"
/misc_difference
21. 22
/gene="At2g20610"
/note="compared to genomic sequence"

```

---

```

/replaces="ttc"
88. 1476
/gene="At2g20610"
/codon_start=1
/evidence=experimental
/product="putative tyrosine aminotransferase"
/protein_id="AAK93664.1"
/db_xref="GI:15293107"
/translation="MSEQPHANLAVPAFKTEKPTQTRNGSSVVRFGGSKAAGA
STVTRAGVIMFLDNCQKVNKILPLGHGDSVYPCFRTCTEARDVAVDLRSGKGN
SYGPGAILPARAVADYNNRDLPHKLTPEIDFLTAGCNGIEIVFESLARPNANTILL
PRGPGHYDARAAYSGLVRYKEDLLPEKEWEIDLEGIEIAEASNTVAVVTPNPFCCG
NVYSHDLKVAETARKLIMWISDEVIRDTIFGDNFFVSMGKFAFIVPVIAGLSK
FWVPGWKIAGIALNDPGEVFFETTKVLSIKONLDVTPDPATIIQAALPALEKADKN
FFAKNKILKHNDVLYCDRLKQIPCVCPKPSCTYLLTLELSLMNDIKDIDFCV
KLAREENLVFLPDGALGNMWRITITGEVHMLDEALERLKGFCFTHAKTKTETETSL
QALKLSDNNLEM"
1477. 1683
3'UTR
1477. 1683
/gene="At2g20610"
misc_difference 1667
/gene="At2g20610"
/note="not present in genomic sequence"
BASE COUNT 497 a 356 c 384 g 446 t
ORIGIN
Alignment Scores:
Pred. No.: 1,68e-71 Length: 1683
Score: 124.00 Matches: 211
Percent Similarity: 66.52% Conservative: 87
Best Local Similarity: 47.10% Mismatches: 138
Query Match: 47.03% Indels: 12
DB: 8 Gaps: 4
US-10-019-783-1 (1-461) x AY050987 (1-1683)
QY 8 HisGlyGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 27
Db |||||
QY 28 AsnGlyLysSerAsnGlyHisAlaAlaAlaAlaAlaAlaValGluTrpAsnPhelaargGly 47
Db |||||
QY 166 AATGGTCAAGTAGC-----GTTGGCGTTTCGGT---GGA 198
QY 48 LysAspGlyLeuAlaAlaThrThrGlyAlaLysAsnSerIleArgAlaIleargTyLys 67
Db |||||
QY 199 AGTGAT-----AAGCAGCAGGAAGCATCCACCGTAACGTTAGAGGTCTCATCATG 252
QY 68 IleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAsp 87
Db |||||
QY 253 CTCTTCGACAACTGCGCGAAAGACGCTCAATAAGACCAATTTTACCCCTCGCCACGGTGAC 312
QY 88 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 107
Db |||||
QY 313 CTTTCGCTTACCTTGGTTCCTGCTACCTGTATCGAAGCTGAAGACCGCTGCTCGACGTC 372
QY 108 LeuArgThrGlyGlnPheAsnCysTyAlaAlaGlyValGlyLeuProAlaAlaArgSer 127
Db |||||
QY 373 CTTGCTCGCGAAAGGCAATCTTACGCTCCGAGCTGGGATCTCCGCCAAGACGA 432
QY 128 AlaValAlaGluHisLeuSerGlnGlyValProTyLysLeuSerAlaAspValPhe 147
Db |||||
QY 433 GCGTTGCTGTATATATGAACCGAGATCTTCGCAACAGTTAAACGCTGAATATTTT 492
QY 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAla 167
Db |||||
QY 493 CTGACCGCTGGATGCAACCAAGGATAGAGATCGGTTCGATCTGTTGGCTCGA---CCA 549
QY 168 GlyAlaAsnIleLeuLeuProArgProGlyTyProAsnTyArgAlaAlaAlaPhe 187
Db |||||
QY 550 AACGCAACATCTTCTCCACGCTCGCTTCCCTCACTACGACGCTGCTGCTGTAC 609
QY 188 AsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIle 207
Db |||||

```

```
Db 610 AGTGGTCTCGAGGTTCCGAAGTTTGATCTTCTTCCCGAGAAGAAATCGGAGATTGATCTT 669
Qy 208 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 670 GAAGTATCGAAGCCATTGCGAGCGAAGACACTGTGGCTATGGTTGTAATTAACCCCAAC 729
Qy 228 AsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArg 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 730 AATCCCTGTGGAATGTTCTACTCTCAGACCATCTCAAAAAGGTTGCGAGAGCGGTAGG 789
Qy 248 LysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 790 AAGCTCGGGATAATGGTGATCTCAGACCAAGTATATGACCGAATATATTCGGAGACAAT 849
Qy 268 ProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeu 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 850 CCAITTTGTTCAATGGGAAGTTTGCTTCGATAGTCCCTGTATTGACACTAGCAGGCATA 909
Qy 288 SerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProThr 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 910 TCTAAGGGATGGGTGTTCTCTGGATGAAAAATTCGCTGGATCGCTTGAATGATCCCGAG 969
Qy 308 LysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 970 GCGCTTTTCGAGACACCAAGGTGTTACAATCCATCAACAGAAATCTTGACGTAACCTCT 1029
Qy 328 AspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAsp 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1030 GACCTGCGCACAAATATTCAGCTGCACTTCACCGATCCTGGAGAAAGCGGACAAAAAC 1089
Qy 348 PhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIle 367
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1090 TTCTTTGCAAGAAGAACAGATATCTCAACATATCTTGATTGTTGTTGTGTGTAGGCTC 1149
Qy 368 LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 387
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1150 AAGACATCCCTCGTGTCTGTCTCCCAAGAACTGAGCTTCGACTTACTTATTGACA 1209
Qy 388 LysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysLeu 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1210 AAGTTGGAGCTGTCTATTGATGATAATATCAAGGACGATATAGATTTTTCGTAATAACTG 1269
Qy 408 AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpVal 427
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1270 GCCAGAGAGGAGAACTCTGTGTTTCTACAGGGGATGCTCTGGGTTTGAAGAACTGGATG 1329
Qy 428 ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSer 447
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1330 AGGATAACCATCGGAGTCGAAGCTCATATGCTTGAGGATGCACCTTGAGAGACTGAGGGT 1389
Qy 448 PheCysGlnArgAsnLysLys 455
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1390 TTCTGTACACGTCAATGCCAAGAAG 1413
```

Search completed: November 8, 2003, 23:37:49  
Job time : 5155.74 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 20:02:05 ; Search time 78.3518 Seconds  
(without alignments)  
2596.975 Million cell updates/sec

Title: US-10-019-783-1

Perfect score: 2390

Sequence: 1 WHQNGHGEAAAAAANKS.....LKVKSFCQRKKNSINGC 461

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 segs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-O=/cgn2\_1/USPTO/spool/US10019783/runat\_07112003\_161644\_19870/app\_query.fasta\_1.1358

-RB=Issued\_Patents\_NA -QFMR=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10019783 @CGN 1 1.76 @runat\_07112003\_161644\_19870 -NCPU=6 -ICPU=3

-NM MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/6C.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	361.5	15.1	1185	2 US-08-646-590B-39
2	361.5	15.1	1185	3 US-09-412-184-39
3	361	15.1	4403765	3 US-09-103-840A-2
4	361	15.1	4411529	3 US-09-103-840A-2
5	328.5	13.7	1527	4 US-09-252-991A-9583
6	325	13.6	1677	4 US-09-252-991A-9544
7	319.5	13.4	7939	4 US-08-961-527-9
8	319.5	13.3	1482	4 US-09-328-352-399
9	315.5	13.2	1236	4 US-09-107-532A-1288
10	315	13.2	1664976	4 US-08-916-421B-1
11	313	13.1	7818	4 US-09-634-238-14
12	308.5	12.9	1122	1 US-08-599-171A-18

13	308.5	12.9	1122	2 US-08-646-590B-18
14	308.5	12.9	1122	3 US-09-069-326-18
15	308.5	12.9	1122	3 US-09-412-184-18
16	305	12.8	1173	4 US-09-107-532A-1579
17	301	12.6	1830121	4 US-09-557-884-1
18	301	12.6	1830121	4 US-09-643-990A-1
19	299	12.5	981	4 US-09-252-991A-9707
20	295.5	12.4	1170	4 US-09-724-623-5
21	291.5	11.8	1107	4 US-09-724-623-4
22	276	11.5	1032	4 US-09-107-532A-1892
23	268.5	11.2	1257	4 US-09-328-352-399
24	265.5	11.1	1197	1 US-08-599-171A-21
25	265.5	11.1	1197	2 US-08-646-590B-21
26	265.5	11.1	1197	3 US-09-069-226-21
27	265.5	11.1	1197	3 US-09-412-184-21
28	260	10.9	1206	4 US-09-252-991A-10627
29	251	10.5	1149	4 US-09-107-532A-2417
30	247	10.3	28473	4 US-08-961-527-83
31	243	10.2	1975	3 US-09-109-204-4
32	243	10.2	1975	4 US-09-490-032-4
33	236	9.9	1233	4 US-09-252-991A-16539
34	236	9.9	1770	4 US-09-252-991A-16060
35	235.5	9.9	1664976	4 US-08-916-421B-1
36	233.5	9.8	4411529	3 US-09-103-840A-1
37	228.5	9.5	4403765	3 US-09-103-840A-2
38	227.5	9.5	1323	4 US-09-252-991A-6342
39	227	9.5	1197	4 US-09-352-991A-6207
40	221	9.2	1701	3 US-08-599-968-2
41	213	8.9	1505	4 US-09-142-481-1
42	212	8.9	1491	2 US-08-941-647A-4
43	210	8.8	1748	3 US-08-765-889C-1
44	210	8.8	1748	5 PCT-US85-07855-1
45	209.5	8.8	1497	4 US-09-328-352-880

#### ALIGNMENTS

RESULT 1

US-08-646-590B-39

; Sequence 39, Application US/08646590B

; Patent No. 5962283

; GENERAL INFORMATION:

; APPLICANT: Warren, Patrick V.

; APPLICANT: Swanson, Ronald V.

; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson, P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/646,590B

; FILING DATE: 08-May-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/599,171

; FILING DATE: 09-FEB-1996

; PRIOR APPLICATION DATA: PCT/US97/01094

; APPLICATION NUMBER: PCT/US97/01094

; FILING DATE: 21-January-1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Haile, Ph.D., Lisa A.

; REGISTRATION NUMBER: 38,347

; REFERENCE/DOCKET NUMBER: 09010/017001

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070  
 TELEFAX: 619/678-5099  
 INFORMATION FOR SEQ ID NO: 39:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1185 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: Genomic DNA  
 FEATURE:  
 NAME/KEY: Coding Sequence  
 LOCATION: 1...1182  
 US-08-646-590B-39

## Alignment Scores:

Pred. No.: 1,41e-30 Length: 1185  
 Score: 361.50 Matches: 109  
 Percent Similarity: 46.57% Conservative: 81  
 Best Local Similarity: 26.72% Mismatches: 173  
 Query Match: 15.13% Indels: 45  
 DB: 2 Gaps: 14

US-10-019-783-1 (1-461) x US-08-646-590B-39 (1-1185)

```

QY 64 lleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeuProLeu 83
Db 1 ATGAGAAAAGGACTTGCAGTAGGTAAGTCAACCTAAACCTCCCCACGCTGACCAT 60
QY 84 AlaHis-----GlyAspProSerVal 90
Db 61 ACCGCAAAAGCAAAAGCAATTAAAGGCTAAAGAGTGGAGCTTAGGTTTGGAGCGGA 120
QY 91 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaLeuArgThr 110
Db 121 GAACCTGACTTCGACACACCGGACTTCATAAAGGAAGGCTGTATAAGGGCTTTAAGGAA 180
QY 111 GlyGlnPheAsnGlyTyrAlaAlaGlyValGlyLeuProAlaAlaAtcSerAlaValAla 130
Db 181 GGAAGAACCAAG---TAGCTCCCTCCCGGGGAATACCGAGGCTCAGAGAGCTATAGCT 237
QY 131 GluHisLeu-----SerGlnGlyValProTyrLysLeuSerAlaAspAspValPheLeu 148
Db 238 GAAAACTACTGAAAGAAAACAAAGTTGAGTACAAACCTTCA-----GAGATAGTCGT 291
QY 149 ThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGly 168
Db 292 TCCGAGGAGCGGAAATGGTTCTCTCTCATATTTCATGGCTATATCTGGAC---GAAGGA 348
QY 169 AlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsn 188
Db 349 GACGAGGTTTACTACTACCTAGCCCTTACTGGTAACTTACCCGAA----- 393
QY 189 LysLeuGluValArgHisPheAspLeuPro-----AspLys 201
Db 394 -----CAGATAAGGTTCTTCGAGGGGTTCCGTTGAGGTTCCCTCTAAAGAAAGAGAA 447
QY 202 GlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMet 221
Db 448 GGATTTCAATTAACTGGAAGATGTGAAGAAAAGGTTACGGAGAGACAAAGCTATA 507
QY 222 ValIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLys 241
Db 508 GTCATAAACTCTCCGAAACAAACCCACTGGTGTGTTTACGAAGAGAGGAACTTAAGAA 567
QY 242 ValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLys 261
Db 568 ATAGCGGAGTTTGGTGGAGAGGGGCAATTTCAATTTCCGATGAGTGTATGAGTAC 627
QY 262 LeuValLeuGlySerAlaProPheLeuProMetGlyValPheGlyHisIleAlaPro--- 280
Db 628 TTCGTTTACGGTGTATGCAAAATTTGTAGCCCTGCTCTTCTCGGATGAAGTAAAGAAC 687

```

```

QY 281 ---ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGly 299
Db 688 ATAACCTTCACGGTAAAGCCCTTTTCGAAGAGCTATTCATGACTGGTTCGGAATAGT 747
QY 300 TrpValAlaValTyrAspProThrLysIleLeuGluLysThrLysIleSerThrSerIle 319
Db 748 TATGTAGCGTCCCGCAA-----GAGTAGCGCAAAAGTATAGCGAGTCTT 792
QY 320 ThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLys 339
Db 793 ---AACAGCCAGAGTGTTC---AACGTCACTACCTTTGCCCATATGAGCTCTTGAG 846
QY 340 IleLeuGluAsnThrLysAla---AspPhePheLysArgIleIleGlyLeuLeuLysGlu 358
Db 847 GCCTTGAAAATCCAAAGTCTAAAGATTTGTAAACGAATGAGAAATGCTTTTGAAGG 906
QY 359 SerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLys 378
Db 907 AGAAGGATACGGCTGTAGAAGAGCTTTCTAAATTCACAGGTATGATGTGTA---AAA 963
QY 379 ProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGluLeuHis 398
Db 964 CCGGAGGTGCCTTTTACATATTTCGGACTTCTCCGCTTAC---GCTGAGAAACTGGGT 1020
QY 399 AspAspIleAspPheCysLysLysLeuAlaLysGluGluSerValIleLeuCysProGly 418
Db 1021 GGTGATGCTAAACTTCGGAGTCTCTCTGGAAGGCTAAGTTTCGGTGGTTCCTCGGT 1080
QY 419 SerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeu 438
Db 1081 TCGGCTTCGGAGCTCCCGGATTTTGGAGCTTCTTACGCCCTTTCCGAGGAAAGACTC 1140
QY 439 GlnAspGlyLeuGluArgValLys 446
Db 1141 GTTAGGGTATAAGGAGAAATAAG 1164

RESULT 2
US-09-412-184-39
; Sequence 39, Application US/09412184
; Patent No. 6286188
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412,184
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,590
; FILING DATE: 08-May-1996
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347

```

REFERENCE/DOCKET NUMBER: 09010/017001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 1185 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 1...1182

US-09-412-184-39

Alignment Scores:

Pred. No.:	1,41e-30	Length:	1185
Score:	361.50	Matches:	109
Percent Similarity:	46.57%	Conservative:	81
Best Local Similarity:	26.72%	Mismatches:	173
Query Match:	15.13%	Indels:	45
DB:	3	Gaps:	14

US-10-019-783-1 (1-461) x US-09-412-184-39 (1-1185)

```

QY 64 ileArgTyrLysileSerAlaSerValGluGluSerGlyProArgProValLeuProLeu 83
Db 1 ATGAGAAAGGACTTGCAGTAGGTAGTCACTTAAACCTTCCCCACGCTGACCATTA 60
QY 84 AlaHis-----GlyAspProSerVal 90
Db 61 ACCGCAAAAGCAAAAGTAAGGGCTAAAGGAGTGGAGCTTATAGGTTTGGAGCGGGA 120
QY 91 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaAlaArgThr 110
Db 121 GAACCTGACTTCGACACACCCGACTTATAGGAGGCTGTATAGGCTTTTAAAGGAA 180
QY 111 GlyGlnPheAsnCyTyAlaAlaGlyValGlyLeuProAlaAlaAArgSerAlaValAla 130
Db 181 GGAAAGACCAAG---TAGCTCCCTCCGGGGAATACCAGAGCTCAGAGAGCTATAGCT 237
QY 131 GluHisLeu-----SerGlnGlyValProTyrLysLeuSerAlaAspValPheLeu 148
Db 238 GAAAACTACTGAAGAAACAAAGTGGATGACAAACCTTCA-----GAGATAGTCGTT 291
QY 149 ThrAlaGlyGlyThrGluAlaGluVallelleProValLeuAlaGlnThrAlaGly 168
Db 292 TCCGAGGAGCGAAATGCTTCTCTCTCATATTATGCTGCTTACTGAC---GAAGGA 348
QY 169 AlaasnlleLeuLeuProArgProGlyTyrProasntyrGluAlaArgAlaAlaPheAsn 188
Db 349 GACGAGGTTTACTACCTAGGCTTACTGGTAACTTACCCCGAA----- 393
QY 189 LysLeuGluValArgHisPheAspLeullePro-----AspLys 201
Db 394 -----CAGATAAGTTCTTCGAGGGGTTCCTGAGGTTCCTCTTAAAGAAAGAGAA 447
QY 202 GlyTrpGluileAspIleAspSerLeuGluSerileAlaAspLysAsnThrThrAlaMet 221
Db 448 GCAATTCATTAAGTCTGGAAGATGTGAAGAAAGAGTTACGGAGAGCAACAAAGCTATA 507
QY 222 ValilleleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLys 241
Db 508 GTCAATAACTCTCCGAAACACCCCTGCTGCTGTATTACGAAGAGAGGAACTTAAAGAA 567
QY 242 VallAlaGluValAlaArgLysLeuGlyileLeuValilleAlaAspGluValTyrGlyLys 261
Db 568 ATAGCGGAGTTTGGGTGGAGAGGGGCATTTTCATAATTTCCGATGAGTGTATGAGTAC 627
QY 262 LeuValleLeuGlySerAlaProPheIleProMetGlyValPheGlyHisileAlaPro--- 280
Db 628 TTCGTTTACGGGTGATGCAAAATTTGTTAGCCCTGCTCTTTCTCGGATGAAGTAAGAAC 687

```

```

QY 281 ---ValLeuSerileGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGly 299
Db 688 ATAACTTTCAGGTAAACGGCTTTTTCGAGAGCTATTCCATGCTGCTGGCAATAGGT 747
QY 300 TrpValAlaValTyrAspProThrLysileLeuGluLysThrLysileSerThrSerile 319
Db 748 TATGTAGCGTCCCGGAA-----GAGTACGCAAAAGTATAGCGAGTCTT 792
QY 320 ThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLys 339
Db 793 ---AACAGCCAGAGTGTTC---AACGTCACCTTTGCCAGTAGTAGGCTCTTGAG 846
QY 340 IleLeuGluAsnThrLysAla---AspPhePheLysArgilelleGlyLeuLeuLysGlu 358
Db 847 GCCTTGAATAATCAAAAGTCTAAAGATTTTGTAAACGAAATGAGAAATGCTTTTGAAGG 906
QY 359 SerSerGluileCysTyrArgGluileLysGluAsnLysTyrileThrCysProHisLys 378
Db 907 AGAAGGATACGGCTGTAGAGAGCTTCTTAAATTCAGAGTATGAGTGTGTA---AAA 963
QY 379 ProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGluileHis 398
Db 964 CCGAGGTGCTTTTACATATTTCGGAGCTTCTCGGCTTAC---GCTGAGAAACTGGGT 1020
QY 399 AspAspIleAspPheCysLysLysLeuAlaLysGluLysSerValilleLeuCysProGly 418
Db 1021 GGTCATGTGAACCTCCGAGTTCCTTCGAAAGAGCTAAGGTTCGGGTGCTTCCCGGT 1080
QY 419 SerValLeuGlyMetGluAsnTrpValArgileThrPheAlaCysValProSerSerLeu 438
Db 1081 TCGGCTTCGAGCTCCGGATTTTGAGGCTTTTACGCCCTTTTACGCCCTTCCGAGGAAAGACTC 1140
QY 439 GlnAspGlyLeuGluArgValLys 446
Db 1141 GTTGGGTTATAAGGAGATAAAG 1164

```

# RESULT 3

US-09-103-840A-2/c  
 ; Sequence 2, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.  
 ; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; FILE REFERENCE: 24366-20007.00  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 4403765  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; FEATURE:  
 ; OTHER INFORMATION: CDC 1551  
 ; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
 ; OTHER INFORMATION: represent a, t, c or g  
 US-09-103-840A-2

Alignment Scores:	1.34e-24	Length:	4403765
Pred. No.:	361.00	Matches:	132
Score:	45.24%	Conservative:	82
Best Local Similarity:	27.91%	Mismatches:	204
Query Match:	15.10%	Indels:	55
DB:	3	Gaps:	18

US-10-019-783-1 (1-461) x US-09-103-840A-2 (1-4403765)

QY 9 GlyGluAlaAlaAlaAlaAlaAsnGlyLysSerAsnGly----- 22  
Db 403297 GGTGAAGGGCTGGGCAATTCAGTGGCGGGCGCGTCCACCGGGCAACCGTGATCACGATT 403238  
QY 23 ---HisAlaAlaAlaAlaAsn-----GlyLysSerAsnGly 33  
Db 403237 TGGACACCTGACAGCAAAATTCAGTGGGACACGATGGACACATTGGACGTGACAAACCA 403178  
QY 34 HisAlaAlaAlaAlaAlaValGluTrpAsnPheAlaArgGlyLysAspGlyIleLeuAla 53  
Db 403177 CCAGCTGCCTGGCACACCGCGCAGCATCAGCGGCGCGCTTCGCTCAGTGGCCAA 403118  
QY 54 ThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSerAlaSerValG1 73  
Db 403117 GCTGACGAGCTCTGTGTACGAGATCCGCGCGCGTGCACCGACGCGCGCGCTGCA 403058  
QY 73 uGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAl 93  
Db 403057 AGCCGAAGGTCAACGC---ATCCTCAAACTCAACATCGCAACCGCGCGCTTC--- 403006  
QY 93 pheArgThrAlaValGluAlaGluAspAla-----ValAlaAlaAlaLeuAr 109  
Db 403005 -----GGCTTCGAAGCACCGCAGCTGATCATCGCGGATATCATCCAGCGCGCTGCC 402956  
QY 109 gThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaVa 129  
Db 402955 CTACCGCGCAG-----GGGTACTCGACTCGACGGCATCTGTGCGCGCGCGTGGGT 402902  
QY 129 lAlaGlu-----HisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPh 147  
Db 402901 GGTACGCGCTACGAGCTGGTGGCGGATTTCCC---CGATTCCAGCTGGACGACGCTA 402845  
QY 147 eLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAl 167  
Db 402844 CTGGGTAAAGGGTCTCGAGCTGATCAGATGAGCTGCAAGCCCTGCTGGACAC-- 402787  
QY 167 aGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPh 187  
Db 402786 -GGCGATCAGGTGCTGATTTCATCCACCGGACTATCCCGCTGTGGACGGCGTGCACCTCGCT 402728  
QY 187 eAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspI1 207  
Db 402727 GCTGGCGGCACTCCGCTCCACTCTCTGTGATGAGACCAAGCTGGACCGCGATAT 402668  
QY 207 eAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAs 227  
Db 402667 CCGCGACCTGGAATCCAGATCACCAGCGCAGCAGCGCTGTGTGATCAACCCCAA 402608  
QY 227 nAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaAr 247  
Db 402607 CAACCAACCGCGGGGTACAGTGGGAATCTCACCAGATGTCGATCTGGCGCG 402548  
QY 247 gLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAl 267  
Db 402547 CAAGCATCACTGCTGCTTGGCGGAGCAATCTACGACAAATCTCTACGACGCGC 402488  
QY 267 aProPheIleProMetGlyValPheGlyHisIleAlaPro-----ValLeuSerI1 284  
Db 402487 CAAGCACATCAACCTG-----GCATCATGATCGCCCGGATATGTGTCCGTGACCTT 402437  
QY 284 eGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTy 304  
Db 402436 CAATGGCTGTGCAAGGCTACCGCTGCCGATACCGCGCGCTGTGCTGGCGATCAC 402377  
QY 304 rAspProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAs 324  
Db 402376 CGGACCCCAAGGAG-----CAGCGCAGCAGCTTTCATCGAGGCGCATCGCGCTGGCCAA 402323  
QY 324 nValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnTh 344  
Db 402322 TATGGTGTGTCCTCAATGTCCCGGCCAGATGCCATTCAGGTGACCTGGCGCGCA 402263  
QY 344 rLysAlaAspPhePheLysArgIleIle-----GlyLeuLeuLysGluSerSerGl 361

Db 402262 TCAGAC-----ATCGAGGACCTGTGTCTGCCCGCGCGCTGCTCGAGCAGCGCA 402209  
QY 361 uLeuCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluG1 381  
Db 402208 CATCGCTGGACCAAGCTCAACGAGATCCCGGGGGTGTGTGCTGTC---AAACCGCGCGG 402152  
QY 381 ySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleHisAspAspI1 401  
Db 402151 CGCGCTGTATGGCTTCCCGGCTAGACCCCGAGGTCTAC---GACATCGAGCAGCACA 402095  
QY 401 eAspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerValIe 421  
Db 402094 GCAACTCGTGTCTGCTGCTGTGCGAGAAGATCTGTGTCAACCCAGGCGACCGGTT 402035  
QY 421 uGlyMetGluAsnTrp-----ValArgIleThrPheAlaCysValProSe 436  
Db 402034 C-----AACTGGCGCGCACCGGATCATCTGCGCTGTGTGACGTCCTCATGTCCCG 401984  
QY 436 rSerLeuGlnAspGlyLeuGluArgValLysSerPhe 448  
Db 401983 CGATCTGGCGCGCGCATCGACGCGCTGGTAACTTC 401947  
RESULT 4  
US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37RV  
US-09-103-840A-1  
Alignment Scores:  
Pred. No.: 1,34e-24 Length: 4411529  
Score: 361.00 Matches: 132  
Percent Similarity: 45.24% Conservative: 82  
Best Local Similarity: 27.91% Mismatches: 204  
Query Match: 15.10% Indels: 55  
DB: 3 Gaps: 18  
US-10-019-783-1 (1-461) x US-09-103-840A-1 (1-4411529)  
QY 9 GlyGluAlaAlaAlaAlaAlaAsnGlyLysSerAsnGly----- 22  
Db 403242 GGTGAAGGGCTGGGCAATTCAGTGGGCGCGCTCCACCGGGCAACCGTGATCACGATT 403183  
QY 23 ---HisAlaAlaAlaAlaAsn-----GlyLysSerAsnGly 33  
Db 403182 TGSCACCTGACAGCAAAATTCAGTGGGACAAACGATGGACACCATGTGTGGAGTGACAAACCA 403123  
QY 34 HisAlaAlaAlaAlaAlaValGluTrpAsnPheAlaArgGlyLysAspGlyIleLeuAla 53  
Db 403122 CCAGCTGCCTGGCACACCGCGCAGCATCAGCGGCGAGCGCGCTTCGCTCAGTCCGCCAA 403063  
QY 54 ThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSerAlaSerValG1 73  
Db 403062 GCTGACGAGCGCTCTGTACGAGATCCGCGCGCGTGTGACCGACGCGCGCGGTCCA 403003  
QY 73 uGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAl 93

Db 403002 AGCCGAGTCCACCG---ATCCTCAAACTCAACATCGACCAACCGCGCGCTTC----- 402951  
Qy 93 aPheArgThrAlaValGluAlaGluAspAla-----ValAlaAlaLeuAr 109  
Db 402950 -----GGCTTCGAAGCACCAGCGATGATCATGCGCGATATCATCCAGCGCGCTGCC 402901  
Qy 109 gThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaVa 129  
Db 402900 CTACGCGCAG-----GGGTACTCCGACTCGCAGGCGATCTCTGTCGCGCGCGCGT 402847  
Qy 129 lAlaGlu-----HisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValph 147  
Db 402846 GGTACGCGCTACGAGTGTGTCGCGATTTCC---CGATTGCGAGTGGAGCGTCTA 402790  
Qy 147 eLeuThrAlaGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAl 167  
Db 402789 CCTGGGTAAACGGGTCTCCGAGCTGATCAGCATGACGCTGCAAGCCCTGCTGGCAAC-- 402732  
Qy 167 aglyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaph 187  
Db 402731 -GGCATCAGGTGCTGATTCATCACCAGACTACCCGCTGTGGAGCGCGCTGACCTGCT 402673  
Qy 187 eAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspI 207  
Db 402672 GGTGGGCGCACTCCGCTCCACTACTGTGCGATGAGACCAAGCTGGAGCCCGATAT 402613  
Qy 207 eAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAs 227  
Db 402612 CCGCGACTGGGAATCAAGATCACCAGCGCACCAAGCGCTGTGCTGATCAACCCCAA 402553  
Qy 227 nAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaAr 247  
Db 402552 CAAACCAACCGCGCGGTGTACAGTGTGGAATCTCCACCAGATGTGCTGATCTGGCCCG 402493  
Qy 247 glyLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAl 267  
Db 402492 CAAGCATCAACTGCTGCTGTGGCGGACGAAATCTACGACAAATCTCTACGACGCGC 402433  
Qy 267 aProPheIleProMetGlyValPheGlyHisIleAlaPro-----ValLeuSerIl 284  
Db 402432 CAAGCAGATACCGCTG-----GCATCGATCCCGCGGATATGTTGCTGCCGTGACCT 402382  
Qy 284 eGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTy 304  
Db 402381 CAATGGGCTGTGAAGGCTACCGGTGTCGCGATACCGCGCGGTGCTGCTGCGATCAC 402322  
Qy 304 rAspProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAs 324  
Db 402321 CGGACCCCAAGGAG-----CAGCGCCAGCAGCTTCATCGAGGCGCATCGGCTGTGGCCAA 402268  
Qy 324 nValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnTh 344  
Db 402267 TATGGGTGTGCCCAATGTCCCGGCCAGCATGCCATGCCATTCAGTTGCTGCGCGGCCA 402208  
Qy 344 rLysAlaAspPhePheLysArgIleIle-----GlyLeuLeuLysGluSerSergl 361  
Db 402207 TCAGAGC-----ATFCAGGAGCCTGTGCTGCCCGCGCGCGCTGCTCGAGCAGCGCA 402154  
Qy 361 uIleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGl 381  
Db 402153 CATGCCCTGGACCAAGCTCAAGAGATCCCGGGGTGTGCTGCTC---AAACCGCGGG 402097  
Qy 381 ySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspI 401  
Db 402096 CGCGTGTATGCTGTTCCTCCGCTAGACCCCGAGGTCTAC--GACATCGACGACGACA 402040  
Qy 401 eAspPheCysCysLysLeuAlaLysGluLysSerValIleLeuCysProGlySerValle 421  
Db 402039 GCAACTGCTGCTGATCTGCTGTGTGCGGAAAGATCTGTGTACCCAGGCGCACCGGTT 401980  
Qy 421 uGlyMetGluAsnTrp-----ValArgIleThrPheAlaCysValProSe 436

Db 401979 C-----AATGGCCCGCAGCCGATCATCTCGCGCTGCTGACGCTGCCATGCTCCG 401929  
Qy 436 rSerLeuGlnAspGlyLeuGluArgValLysSerPhe 448  
Db 401928 CGATCTGGCGCGCCCATCGAGCGCTGGTAACTTC 401892  
RESULT 5  
US-09-252-991A-9583  
; Sequence 9583, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 9583  
; LENGTH: 1527  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-9583  
Alignment Scores:  
Pred. No.: 1,04e-26 Length: 1527  
Score: 328.50 Matches: 117  
Percent Similarity: 42.76% Conservative: 75  
Best Local Similarity: 26.06% Mismatches: 202  
Query Match: 13.74% Indels: 55  
DB: 4 Gaps: 14  
US-10-019-783-1 (1-461) x US-09-252-991A-9583 (1-1527)  
Qy 27 AlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGluTrpAsnPheAlaArg 46  
Db 244 GCCAAGCGCGATCCAGCGGCTTTTCGCAGCGCTGTTCACGGTG-----TTGCTCAC 237  
Qy 47 GlyLysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyr 66  
Db 298 TTCGGGACTCGACATCATGTCAGGTCAGCAATCGACAAGCTGCCACGCTCTCTAT 357  
Qy 67 LysIleSerAlaSerVal-----GluGluSerGlyProArgPro 79  
Db 358 GACATTCGGCGCGCGTGTCTCAAGCACGCGCAACGCTAGAGGAGAGGCTCACCAGC-- 414  
Qy 80 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 99  
Db 415 ATCTCTCAAGCTGAATATCGCAACCGCGCGCTTC---GGTTTCGAAGCTCCCGAGGAA 471  
Qy 100 AlaGluAspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGly 119  
Db 472 ATTCTCCAGGACGTATCCGCAACCTGCGACCGCCAG-----GGCTACAGCGACTCC 525  
Qy 120 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 139  
Db 526 AAGGCGCTGTTCAGCGCGCGCGAGGAGTATTCAGTATTACGACGAGAGGAGGTGGAA 585  
Qy 140 LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159  
Db 586 GCGGTTGGCATCGAGGACATCTACCTCGGACCGGCTGTCGAACTGATCGTGAIGTCC 645  
Qy 160 IleProValIleAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 179  
Db 646 ATGACGCGCTGTGTGAACAAAC---GGCGACGAGGTGCTGATCCCGCTCCGACTACCG 702  
Qy 180 AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199  
Db 703 CTGTGACCGCGCGTCTAGCTCCCGCGCGGAGCGGCTGCTCTCTGCGGACGAG 762

```
QY 200 AspLysGlyTTPGluIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219
DB 763 CAGGCCAACTGGTGGCGGACTGGAAGACATCAAGCGGAAGATCATCGCCGACACACAGG 822
QY 220 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239
DB 823 GCGATGGTGATCATCAACCCGAAACCCACCCGCGCGGTGTATTCAGGGAAGTCTG 882
QY 240 AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyr 259
DB 883 GAAGCATGGTGGAACTGGCCGCGCAGCACAACTGGTGGTCTTCGACGAGATCTAC 942
QY 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279
DB 943 GACAGATCTCTACGAGCGGCGCTCCACGCTCTCCACCGCTCGTGGCGCGGACGTG 1002
QY 280 ProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrPheGlyLeuGly 299
DB 1003 CTCTGCCCTGACCTTCAACGGGCTGTGCAAAATCTACCGGGTGGCGGCTTCCGCTCCGGC 1062
QY 300 TrpValAlaValTyrAspProThrLys-----IleLeuGluLysThrLysIle 315
DB 1063 TGGGTGGGATCTCCGGGCCCAAGCAGCGGCGCAGAGCTATATCGAAGGTCTCGACATC 1122
QY 316 SerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGlu 335
DB 1123 CTCGCCAATGCGCTGTGCGCAACGTCCCGGCGCAGCAGCG-----ATCCAGACC 1176
QY 336 Ala-----LeuProLysIleLeuGluAsnThr 344
DB 1177 GCGTGGCGGCTACACAGCATCAACGATCTGGTCTCGCGCGC----- 1221
QY 345 LysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 364
DB 1222 -----GGCGCGCTGTGGAGAGCGCAACCGCGCGCTGG 1254
QY 365 ArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGlyLysSerMetPhe 384
DB 1255 GMACTGCTCAAGCATCCCGGCGTACGTGCGTG---AAGCCGATGGCGCGCTGTAC 1311
QY 385 ValMetValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCys 404
DB 1312 GCCTTCCCGCGGATCGACCCGAAAGTCTGC---CGATCCCAACAGCAGAAAAGTCTGTC 1368
QY 405 CysLysLeuAlaLysGluLysSerValIleLeuCysProGlySerValLeuGlyMetGlu 424
DB 1369 CTCGACCTACTGCTCTCGAAAAACTGCTGTATGTCAGGCGACCGCTTC----- 1419
QY 425 AsnTrp-----ValArgIleThrPheAlaCysValProSerSerLeuGln 439
DB 1420 AACTGCGCGTGGCCCGACCACTTCGGGTGTGTCACTCCCTGCCCGCTCGACGACCTGGAG 1479
QY 440 AspGlyLeuLysArgValLysSerPhe 448
DB 1480 CAGGCGATCTCGCATCGCGAGCTTC 1506
```

## RESULT 6

```
US-09-252-991A-9544
; Sequence 9544, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; TITLE OF INVENTION: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
```

```
; SEQ ID NO 9544
; LENGTH: 1677
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9544
```

```
Alignment Scores:
Pred. No.: 36-26 Length: 1677
Score: 325.00 Matches: 109
Percent Similarity: 44.12% Conservative: 71
Best Local Similarity: 26.72% Mismatches: 182
Query Match: 13.60% Indels: 46
DB: 12
Gaps: 2
```

US-10-019-783-1 (1-461) x US-09-252-991A-9544 (1-1677)

```
QY 61 IleArgAlaIleArgTyrLysIleSerAlaSerValGluSerGlyProArgProVal 80
DB 2 ATTCGGGCGCGTCTCAGCACCCCAACGCTAGAGGAAGGGTCAACGC---ATC 58
QY 81 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 100
DB 59 CTCAAGCTGAATATCGCAACCCGCGCGCTTC---GGTTTCGAAGCTCCCGAGGAAT 115
QY 101 GluAspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyVal 120
DB 116 CTCAGGAGCTATCGCAACCTCGACCGCCAG-----GGCTACGAGCTCCCAAG 169
QY 121 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLys 140
DB 170 GGCCTGTTCCGCGCGCAAGGAGTATGATGATGATGATGATGATGATGATGATGATG 229
QY 141 LeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIle 160
DB 230 GTTGGCATCGAGACATCTACCTCGCAACGCGGTGCGAACTGATGATGATGATGATG 289
QY 161 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 180
DB 290 CAGGCGCTCTCAACAAC---GGCGACGAGGTGTGATCCCGCTCCCGACTACCCGCTG 346
QY 181 TyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIleProAsp 200
DB 347 TGGACCGCGCGGTCAAGCTCCCGCGCGCAAGCGGTGCGACTACTCTGTGGACGAGCAG 406
QY 201 LysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrAla 220
DB 407 GCCAACTGTGGCGGACCTGGAGACATCAGGCGAAGATCACCGCAACACCGAGGCG 466
QY 221 MetValIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAla 240
DB 467 ATGGTGATCATCAACCCGCAACCCACCGCGCGGTGTATTCAGGGAAGTGTGGAA 526
QY 241 LysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGly 260
DB 527 GGCATGTCGAATGGCGCGCGCAGCAACCTGGTGTGTTCTCCGACGAGATCTACGAC 586
QY 261 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro 280
DB 587 AAGATCTCTACGAGCGCGCGCTCCACGCTCTCCACCGCTCGCTCGCGCGCGAGCTGTC 646
QY 281 ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTyr 300
DB 647 TGCCTGACCTTCAAGCGCTGTCCAAATCTACCGGGTGGCGGCTTCCGCTCCGGCTGG 706
QY 301 ValAlaValTyrAspProThrLys-----IleLeuGluLysThrLysIleSer 316
DB 707 GTGGCATCTCCGGGCGCAAGCAGCGGCGCAGAGCTATATCGAAGTCTTCACATCTCTC 766
QY 317 ThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAla 336
DB 767 GCAACATCGCGCTGTGCGCCCAACGTCGCGGCGCAGCAGCG-----ATCCAGACCGCC 820
QY 337 -----LeuProLysIleLeuGluAsnThrLys 345
```

Db 821 CTGGCGGCTACACAGCATCAACGATCTGCTCCGCGCCG----- 862  
 QY 346 AlaAspPhePheLysArgIleIleGlyLeuLysGluSerSerGluIleCysTyrArg 365  
 Db 863 -----GGGCGCTGCTGGAGCAGCGCAACCGCGCTGGGAA 898  
 QY 366 GluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 385  
 Db 899 CTGCTCAACGATCCCGCGCTCAGCTCGNG---AAGCCGATGGGCGCGCTGACGCC 955  
 QY 386 MetValLysLeuAsnLeuHisLeuGluGluIleHisAspAspIleAspPheCysCys 405  
 Db 956 TTCCCGCGGATCGACCGCAAGCTCGC---CCGATCCACACGACGAAAGCTTCGCTCTC 1012  
 QY 406 LysLeuAlaLysGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsn 425  
 Db 1013 GACCTACTGCTCTCGGAATACTGCTGATCGTCACGAGCCCGCTTC-----AAC 1063  
 QY 426 Trp-----ValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 440  
 Db 1064 TGGCGGTGGCCGACCACTTCGGGCTGTCACCTGCCCGGTCGACGACCTGGAGCAG 1123  
 QY 441 GlyLeuGluArgValLysSerPhe 448  
 Db 1124 GCGATCCTCGCATCGCAGCTTC 1147

## RESULT 7

US-08-961-527-9/c  
 ; Sequence 9, Application US/08961527  
 ; Patent No. 6420135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 391  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/961,527  
 ; FILING DATE:

; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brookes, A. Anders  
 ; REGISTRATION NUMBER: 36,373  
 ; REFERENCE/DOCKET NUMBER: PB340P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 9:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 7939 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; US-08-961-527-9

## Alignment Scores:

Pred. No.: 1,63e-24 Length: 7939  
 Score: 319.50 Matches: 94  
 Percent Similarity: 48.17% Conservative: 90

Best Local Similarity: 24.61% Mismatches: 166  
 Query Match: 13.37% Indels: 33  
 DB: 4 Gaps: 8  
 US-10-019-783-1 (1-461) x US-08-961-527-9 (1-7939)  
 QY 78 ArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAla 97  
 Db 7162 AGAGATATTTTGTCTCTAACCCTTGGGTGAGCCAGAT-----TTCACTACTCCC 7115  
 QY 98 ValGluAlaGluAspAlaValAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAla 117  
 Db 7114 AAAAATATCCAGATCCGCCATTCGATTCGATTCGAGATGGGCTGCTCTTTTATACA 7055  
 QY 118 AlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyVal 137  
 Db 7054 GTAACCTCAGTCTGCCAGAGCTTAAGCGGGCGGTCAATAGCTACTTTGAGCGCTTTAC 6995  
 QY 138 ProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGlu 157  
 Db 6994 GGTATTTCTTAGCGTCAATCAAGTGACAGTCGCTCGGGAGCCAAAATATTCTCTCTAT 6935  
 QY 158 ValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGly 177  
 Db 6934 ACCTTCTTTATGGCTGTGGTC---AATCCAGGTGATGAAGTCATCATCCCAACCCATAC 6878  
 QY 178 TyrProAsnTyr-GluAlaArgAlaAlaPheAsnLysLeuGluValAlaArgHisPheAspLe 197  
 Db 6877 TGGGTAGCTATGGAGATCAGGTCAAGATGGCAGAGCGCTTCCCGCTTTTGTCTGCT 6818  
 QY 197 uIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAs 217  
 Db 6817 AAG-GAAGACAATCACTTTAAGGTGACCTAGAGCAGATTAGAGCAGCTCGCACTACAA 6759  
 QY 217 nThrThrAlaMetValIleIleAsnProAsnLysProCysGlySerValTyrSerTyrAs 237  
 Db 6758 GACCAAGGTTTGGTGTGTAATTCGCCACTCTAATCCACAGGTATGATTATACACCGCTGA 6699  
 QY 237 pHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGI 257  
 Db 6698 GGAACCTCTGGCAATTGGAACTGGGCTGTGAATAATGATATTCTCATCTAGCAGCA 6639  
 QY 257 uValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHi 277  
 Db 6638 TATCTATGGCGGCTTGGTTTATAATGGTCATGAGTTCAACCGATTCTAGCCATACGGA 6579  
 QY 277 sIle-----AlaProValLeuSerIleGlySerLeuSerLysSerTyrIleValProGI 295  
 Db 6578 AGGATTCGCAAGCAAAACAGTGGTCAATGGTGTGTCTAAACTTATGCCATGACTGG 6519  
 QY 295 YTrpArgLeuGlyTyrValAlaValTyrAspProThrLysIleLeuGluLysThrLysI 315  
 Db 6518 TTGGCGGATTGGTTAT---GCCGTTGGAGAGCAGACATATCGCTGCCATGCTCAAGAT 6462  
 QY 315 eSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheValGlnGI 335  
 Db 6461 TCCAGGTCAAACT-----TCGAATCCGTGACGAGTAGCCCAATA 6420  
 QY 335 uAlaLeuProLysIleLeu-----GluAsnThrLysAlaAspPh 348  
 Db 6419 TGCAGCAGTTGAGGCTCTATCAGGTGAGCAAGATACTGTAGAAGAGCATCGTCAGGCCCT 6360  
 QY 348 ePheLysArgIleIleGlyLeuLysGluSerSerGluIleCysTyrArgGluIleL 368  
 Db 6359 TGAGGAACGCTCTAATACCATCTATCCCTCTCTTCAGAGGTACACAGGATTTGAAGTGGT 6300  
 QY 368 sGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal 388  
 Db 6299 C-----AAACCGCAAGGGCGCTTCTATCTCTCTCCCAA 6267  
 QY 388 sLeuAsnLeuHisLeuGluGluIleHisAspAspIle---AspPheCysCysLysLe 407  
 Db 6266 TGTCAAAAAGGCCATGGAGATGAAGGCTACACGGATGTGACAGACTTTTACAACTCTTAT 6207

QY 407 uAlaYsGluGluSerValIleuLeuCysProGlySerValLeuGlyMetGluAsnT-pVa 427  
Db 6206 CTTAGAGAACCGAAGTGGCTTGGTAACAGAGAGCTGGATTGGACACACAGAAATGT 6147  
QY 427 lArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSe 447  
Db 6146 CGCCTCAGCTATGCGACAGACCTAGACACGCTTAAAGAACGAGCTGCAACGCTTGAAGC 6087  
QY 447 rPhe 448  
Db 6086 ATTT 6083

RESULT 8  
US-09-328-352-2537  
; Sequence 2537, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328.352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 2537  
; LENGTH: 1482  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-2537

Alignment Scores:  
Pred. No.: 1,14e-25 Length: 1482  
Score: 319.00 Matches: 107  
Percent Similarity: 45.34% Conservative: 73  
Best Local Similarity: 26.95% Mismatches: 193  
Query Match: 13.35% Indels: 24  
Gaps: 12

US-10-019-783-1 (1-461) x US-09-328-352-2537 (1-1482)

QY 61 lIleArgAlaIleArgTyLysIleSerAlaSerValGluGluSerGlyProArgProVal 80  
Db 319 ATTCGGACCAAGTATTACGAGCCGCAATGAATGGAGAGCAGACATAAA---ATC 375  
QY 81 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 100  
Db 376 ATTAAGCTGAATATCGGCAACCTGCTCCATTT---GGTTTGAAGCACCACCAAGAAATT 432  
QY 101 GluAspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaGlyVal 120  
Db 433 ATTAATGACGTT-----GCTTTAACTACCAATGCAATGCTGTTATGATTCATAA 486  
QY 121 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyLys 140  
Db 487 GGTATCTTCCGGCAGCTAAAGCAATCTCCAGTACTACCAGCAAAAGGCAATCTTAAT 546  
QY 141 LeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 160  
Db 547 ATGCACGTTAATGAGCTGTATCGTAAATGCTGTAATGCTGTAATGCTGTAATGCTGTA 606  
QY 161 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyProAsn 180  
Db 607 CAAGCCCTACTCGATGAT---GGTGATGAATGCTCATTCGATGCCGAGCTATCCGCTT 663  
QY 181 TyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIleProAsp 200  
Db 664 TGGACAGCTGCGGTTAACTGTCAGCGGCTGACGCCATTCATTATAAGTGTGATGAAGAA 723  
QY 201 LysGlyTTPGluIleAspSerLeuGluSerIleAlaAspLysAsnThrAla 220  
Db 724 AACAGTTGGTATCCGTGATATCCGATATTCGCGATATTAAGTAAATTAATCTTCAAACTCGTGGT 783

QY 221 MetValIleIleAsnProAsnProCysGlySerValTySerTyAspHisLeuAla 240  
Db 784 ATTGTCATTATCAATCCGAAACCAACTGGTTCGGTATATCCACGTCATGTCGTTGAG 843  
QY 241 LysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyGly 260  
Db 844 CAAATTTGGCAGCTTGCAGAAACATGACCTGATTTTATTTGCTGAGCAATTTAGCAC 903  
QY 261 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro 280  
Db 904 AAATCGTTTACGATGGCATTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963  
QY 281 ValLeuSerIleGlySerLeuSerIleValProGlyTrpArgLeuGlyTrp 300  
Db 964 TGTATTTCTTCAATGGGCTATCAAAAGCATATGATTTGCTGCTGCTGCTGCTGCTGCTG 1023  
QY 301 ValAlaValTyLys---AspProThrLys-----IleLeuGluLysThrLysIleSer 316  
Db 1024 ATGGCAATCACAGGCGATAAGCGGTCGACGACGATATATCGAAGTCTCGATATGTTA 1083  
QY 317 ThrSerIleThrAsnTyLysValSerThrAspProAlaThrPheValGlnGluAla 336  
Db 1084 GCTTCAATCGGTTTATGTCAGAACGTCACAGCTCAGTATGCA-----ATTCAACGCGC 1137  
QY 337 LeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeu 356  
Db 1138 CTGTGTGGCTAT-----CAGTCTATTATGATTTGATTCGACGACGAGTGGAGCTT 1188  
QY 357 LysGluSerSerGluIleCysTyArgGluIleLysGluAsnLysTyIleThrCysPro 376  
Db 1189 TATGAACAGCGTAATATCGCATGGGAAATGTTAAATGAAATTCCTGGTGAAGCTGTT 1248  
QY 377 HisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGlu 396  
Db 1249 ---AAACAGAGAGCGCAATGTTTTCGCGCTGTTGACCCCTAACATTTAC---CCA 1302  
QY 397 lIleHisAspAspIleAspPheCysLysLysLeuAlaLysGluGluSerValIleLeuCys 416  
Db 1303 ATCGAAGATGATGAAAGCTTAATGTTAGATTACTTCGTCGTGAAAGTACTATTAGTT 1362  
QY 417 ProGlySerValLeuGlyMetGluAsnTrp-----ValArgIleThrPhe 431  
Db 1363 CAAGGAACAGGTTT-----AACTGGCCACACACGACCATTTCCGTTGTTTTC 1413  
QY 432 AlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPhe 448  
Db 1414 TTACCGGCAGAAACGAATTAGTGAAGCAATGTCGCTGTTAGTCGTTTC 1464

RESULT 9  
US-09-107-532A-1288  
; Sequence 1288, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; City: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER: PC  
; MEDIUM TYPE: CD-ROM ISO9660  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:



APPLICATION NUMBER: 60/085,598  
 FILING DATE: 14 May 1998  
 APPLICATION NUMBER: 60/051571  
 FILING DATE: July 2, 1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ariniello, Pamela Deneke  
 REGISTRATION NUMBER: 40,489  
 REFERENCE/DOCKET NUMBER: GTC-012  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (781)893-5007  
 TELEFAX: (781)893-8277  
 INFORMATION FOR SEQ ID NO: 1288:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1236 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Enterococcus faecium  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (B) LOCATION 1...1236  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1288:

US-09-107-532A-1288

Alignment Scores:  
 Pred. No.: 2, 08e-25 Length: 1236  
 Score: 315.50 Matches: 110  
 Percent Similarity: 45.27% Conservative: 91  
 Best Local Similarity: 24.77% Mismatches: 170  
 Query Match: 13.20% Indels: 74  
 DB: 4 Gaps: 19

US-10-019-783-1 (1-461) x US-09-107-532A-1288 (1-1236)

Qy 41 GluTrpAsnPheAlaArgGlyLysAspGlyLeuAlaThrThrGlyAlaLysAsn--- 59  
 Db 39 GAATGG-----GATGGAGTTATC-----AAAAAGAACGAA 68  
 Qy 60 SerileArgAlaIleArgTyrLysIleSerAlaSerValGlu-----GluSer 75  
 Db 69 AGCATAGAACCATCTGTACT--TTGGCAGTCGGCAAAAGCAAGCGGCTTAAAGCA 126  
 Qy 76 GlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArg 95  
 Db 127 AAAGGAAGAGATGCTCGATTGACTGTGGAGAACCGGAT-----TTTCGG 174  
 Qy 96 ThrAlaValGluAlaGluAspAlaValAlaAlaLeuArgThrGlyGlnPheAsnCys 115  
 Db 175 ACACCAAGAAACATCCAGAGACGACCAATCGGAAGCAATCCGTAACGGAAAAAGCTAGCTAT 234  
 Qy 116 TyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeu----- 133  
 Db 235 TATACCAACAGCAGAGATCCAGACTTCGACAGGCGATCGTTGATATCTGAAGAA 294  
 Qy 134 SerGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThr 153  
 Db 295 AACGATGACTTGTCTATGAATCTTCA-----CAAAACGGTCGTAACAGATGGCGGAAA 348  
 Qy 154 GlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeu 173  
 Db 349 TTGTCTTTGTATACATCTTTTCAACAATTTGGACCGCGAAGAT--GAAGTCATCAT 405  
 Qy 174 ProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArg 193  
 Db 406 CCGTGGCTTATTTGGTTCAGTACGGCGAG-----CAAGTGAA 444  
 Qy 194 HisPheAspLeuLeuPro-----AspLysGlyTyrGluIleAsp 206  
 Db 445 CTGGCAGAAAGCGCTGCTCTTTTGTGAAGGAAAGAAATCTAATAACTGGAAAGTCACT 504

Qy 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 226  
 Db 505 GTTCAACAGTTAGAACACAGCAGCTTTAGGAAAAACAAAGCCGTGATCATCAATGCCA 564  
 Qy 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246  
 Db 565 TCGAATCCGACTGGATGATTTATAGCAAGATGAAGTGCAGCAATTTGGTGAATGGCA 624  
 Qy 247 ArgLysLeuGlyIleLeuValIleAlaAspValTyrGlyLysLeuValLeuGlySer 266  
 Db 625 GTAAAAAATGATATTTTATAGTGTCTGATGATATTTATGTCGTTAGTTTATTAATGGC 684  
 Qy 267 AlaProPheIleProMetGlyValPheGlyHisIleAla-----ProValLeuSerIle 284  
 Db 685 AATGAGTTTACTCCAAATGCACAGATTTAGAGCAATCAAAATCAGACCATCATCAT 744  
 Qy 285 GlySerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyr 304  
 Db 745 AATGGCGTGTCTAAGACCTATGCGATGACACAGCGTGGCGCATCGTTAT---GGGTTGGG 801  
 Qy 305 AspProThrLysIleLeuGlyLysThrLysIleSerThrSerIleThrAsnTyrLeuAsn 324  
 Db 802 AATCCTGAAATATCAATGGGATGATCGCTGTGCTTCCCAATCAACT-----849  
 Qy 325 ValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThr 344  
 Db 850 -----AGCAATCCGACCGCAGTCAGTCATATGACGCTGTGAGCAATTCGAAGGT--- 900  
 Qy 345 LysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 364  
 Db 901 GAGCAAGATACAGTAGAAGAAATGCGGAACGCCCTTTGAAGACGCGCTGAATACGTTATAT 960  
 Qy 365 ArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySer----- 382  
 Db 961 CCCTTGTGGCAGAACTACAGGTGTTCTCTT---GAAAGCCGCAAGAGGCTTTTAT 1017  
 Qy 383 MetPheValMetValLysLeuAsnLeuHisLeuGluGluIleHisAspPheAsp 402  
 Db 1018 CTTTTCCTTAATGTAAAGAAACGCTTCGCATG-----1050  
 Qy 403 PheCysCysLys-----LeuAlaLysGluGluSer 412  
 Db 1051 -----TGTAAGTATGAAATGTCACAAATGGGTAGAGATCTTTTGAAGAAACACAGGG 1104  
 Qy 413 ValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAla 432  
 Db 1105 GTAGCACTTGTGACTGGCGAAGGATTTGGGCACCGAGAAATGTCGCATGAGTTATGCG 1164  
 Qy 433 CysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452  
 Db 1165 ACTGATCGGGCTACGTTAGAAAGCGGTACGCAAGATCAAGCAATTTATTGAG---AGT 1221  
 Qy 453 LysLysLysAsn 456  
 Db 1222 AAAAGTCAAAAT 1233

RESULT 10

US-08-916-421B-1  
 ; Sequence 1, Application US/08916421B  
 ; Patent No. 6503729  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bult et al.  
 ; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ  
 ; Patent No. 6503729  
 ; FILE OF INVENTION: jannaschii  
 ; FILE REFERENCE: PB275  
 ; CURRENT APPLICATION NUMBER: US/08/916,421B  
 ; CURRENT FILING DATE: 1997-08-22  
 ; PRIOR APPLICATION NUMBER: US 60/024,428  
 ; PRIOR FILING DATE: 1996-08-22  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (103998)..(103998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (148948)..(148948)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (163385)..(163385)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (191989)..(191989)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (191995)..(191995)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (231980)..(231980)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (234187)..(234187)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (234220)..(234220)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (234814)..(234814)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (309398)..(309398)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (309418)..(309418)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (312837)..(312837)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (312993)..(312993)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (319226)..(319226)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (559167)..(559167)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (559241)..(559241)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (600992)..(600992)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (622708)..(622708)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (657081)..(657081)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (657203)..(657203)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (67435)..(67435)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (682442)..(682442)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (713652)..(713652)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (741884)..(741884)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (779455)..(779455)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (779676)..(779676)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
```

NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1  
Alignment Scores:  
Pred. No.: 3,676-20 Length: 1664976  
Score: 315.00 Matches: 93  
Percent Similarity: 45.04% Conservative: 84  
Best Local Similarity: 23.66% Mismatches: 170  
Query Match: 13.18% Indels: 46  
DB: 4 Gaps: 13  
US-10-019-783-1 (1-461) x US-08-916-421B-1 (1-1664976)  
Qy 71 SerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 90  
Db 1469039 AATATTAACCATGACGATTAAGGAGATATTTAACTAGCTACATCTGATGATAAT 1469098  
Qy 91 Phe-----ProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAla 106  
Db 1469099 TTAGGAATAGGAGACCGAGATTTTCACACTCCAAAGCATATCATTTGAGGCTCCAAAAGG 1469158  
Qy 107 AlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArg 126  
Db 1469159 GCTTTAGTAGAGGGGAAACTCAC---TACTCTCCAAACAAATGGAATTCAGAGCTTGA 1469215  
Qy 127 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspVal 146  
Db 1469216 GAGGAGATAGCAATTAAGATGATTAACACTGATTAAGCTTGTATGATTAAGCAATTT 1469275  
Qy 147 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThr 166  
Db 1469276 ATTGTACCTGTGGAGCTTACAGCGTAAATGTTATCTATATGATTTGATTTGAC--- 1469332  
Qy 167 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 186  
Db 1469333 AGAGGGATGAGGTTTAAATCCAAATCCATCTTTTGTGCTTTTTCCTACTAACACAG 1469392  
Qy 187 PheAsnLysLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAsp 206  
Db 1469393 TTTGCTGAGGTTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1469446  
Qy 207 IleAspSerLeuGluSerIleAlaAspLysAsnThrAlaMetValIleIleAsnPro 226  
Db 1469447 TTAGAGAAGTTAAGAAATCAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1469506  
Qy 227 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAla 246  
Db 1469507 TCAATCTACTGNAAGTCTATGATTAAGAGACCAATAAAGCTTGGCGAGATTTGCT 1469566  
Qy 247 ArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 266  
Db 1469567 GAGGATTAATCTAATTAATTTGTTTTCAGATGAAGCTTATGATTAAGATTAATCTA 1469626  
Qy 267 AlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySer 286  
Db 1469627 AAGCATTAATCTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1469677  
Qy 287 LeuSerLysSerTrpIleValProGlyTyrPargLeuGlyTyrPValAlaValTyrAspPro 306

Db 1469678 TTCTTAAACCTATGACCATGACATGAGGATGAGGATGATTTGCTGTTTCTGAT--- 1469734  
Qy 307 ThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 326  
Db 1469735 -----GNACTAATAATAGGAGTATGATTAATCAACATATGATTAAGATTCAT 1469782  
Qy 327 -----ThrAspProAlaIleThrPheValGln-----Glu 335  
Db 1469783 CAGTATAGCTTTGATGCTGCTACAACTTTGCTCAATATGGGCAATTAGCAGCTTTAAGA 1469842  
Qy 336 AlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeu 355  
Db 1469843 GGTAGTCAAAAGTGTGTGTAGGATATGTTTAGAGATTT----- 1469881  
Qy 356 LeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIleThrCys 375  
Db 1469882 ---AAATGAGAGAGATTTAATCTATATGATTAAGGAT-----AICTTTAA 1469929  
Qy 376 ProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGlu 395  
Db 1469930 GTCAATTAAGCAGATGGGCAATCTTATATATCCAGATGTTTCT----- 1469974  
Qy 396 GluIleHisAspIleAspPheCysCysLysLeuAlaLysGluLysValIleLeu 415  
Db 1469975 GAGTATGAGATGGGTAGAGTAGCTAAGAAATGATT---CAGATTAAGGTTTATGT 1470031  
Qy 416 CysProGlySerValLeuGlyMetGlu-----AsnTrpValArgIleThrPheAlaCys 433  
Db 1470032 GTTCTCGAGCTTCATTTGCTGAGATGAGCTTAACATATATTAGTTTACGCTATGCTACN 1470091  
Qy 434 ValProSerSerLeuGlnAspGlyLeuGluArgValLys 446  
Db 1470092 AATATGAGATATAGAGAAGCTTTAGGATTAATAA 1470130  
RESULT 11  
US-09-634-238-14/c  
; Sequence 14, Application US/09634238  
; Patent No. 6544772  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Havikkala, Ilkka J.  
; APPLICANT: Bloksberg, Leonard, N.  
; APPLICANT: Lubbers, Mark W.  
; APPLICANT: Dekker, James  
; APPLICANT: Christenson, Anna C.  
; APPLICANT: Holland, Ross  
; APPLICANT: O'Toole, Paul W.  
; APPLICANT: Reid, Julian R.  
; APPLICANT: Coolbear, Timothy  
; TITLE OF INVENTION: Polynucleotides, materials incorporating  
; FILE REFERENCE: 11000.1043U1  
; CURRENT APPLICATION NUMBER: US/09/634,238  
; NUMBER OF SEQ ID NOS: 422  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 7818  
; TYPE: DNA  
; ORGANISM: Lactobacillus rhamnosus  
US-09-634-238-14  
Alignment Scores:  
Pred. No.: 8,456-24 Length: 7818  
Score: 313.00 Matches: 113  
Percent Similarity: 42.38% Conservative: 79  
Best Local Similarity: 24.94% Mismatches: 206  
Query Match: 13.10% Indels: 55  
DB: 4 Gaps: 16  
US-10-019-783-1 (1-461) x US-09-634-238-14 (1-7818)  
Qy 25 AlaAlaAlaAsn-----GlyLysSer 31

Db 6961 CGAGCGGCAACATTGACCGCATCAAAATTTTATTAATTAATGATTCGGTAGCCGC 6902  
QY 32 AsnGlyHisAlaAlaAlaValGluTrpAsnPheAlaArgGlyLysAspGlyLe 51  
Db 6901 AATGTCGGAAGCTGTAGGAGGCGCAGTCGTGATTTAGCAAAAGAACCTCTCAACGTC 6942  
QY 52 Leu-----AlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSer 69  
Db 6841 GCACCGTCAGCAGACATTGCGCTTAAGTAATCAGACGAAGACTTAAGGCAAAA----- 6788  
QY 70 AlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSer 89  
Db 6787 -----GGTCCGACGCTCATTTGATTTGCTATTGTCGCAACACAGAT 6749  
QY 90 ValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaLeuArg 109  
Db 6748 -----TTTTCACCCCTTAAGCGGATTGATGACGAGCTATTGCGCGCATTCAG 6701  
QY 110 ThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaVal 129  
Db 6700 GCTGTAATGCCAGTTCTATACGCGCAGCAACCGGTATTCGGAATTAAGCAGCGCAT 6641  
QY 130 AlaGluHisLeu-----SerGlnGlyValProTyrLysLeuSerAlaAspValPhe 147  
Db 6640 AGTGACCGGATATTTGCCAAGACGCTATTCGTTAT-----GATCATCGTCAAAATCGT 6587  
QY 148 LeuThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThrAla 167  
Db 6586 GCAACACCGCGCTAAGTTGCTTGTATGCTTATTCAGGTTTCTTA--AAACCA 6530  
QY 168 GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 187  
Db 6529 GCGGATGAGTGCTGATTCCTGCTCCTACTGCTGTTTCTACGAGGAACAGATT----- 6476  
QY 188 AsnLysLeuGlu-----ValArgHisPheAspLeuIleProAspLysGlyTrpGluLe 205  
Db 6475 ---AAATTGGCGAGCGCGTGCACATCTG--GTATGCGCGGAGCTCGGACATAAAGTC 6422  
QY 206 AspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsn 225  
Db 6421 AGTGTCGATGATCTTGAGCGCGCTCGGACCGCATAAACCGCGGATTCGATTAATTCG 6362  
QY 226 ProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluVal 245  
Db 6361 CCACAAAACCCCAAGTGGCGTGTCTATGATCGCAGCACTGACCTTAATTGGCAATTGG 6302  
QY 246 AlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGly 265  
Db 6301 GCGCTGAACCATCATATTTGGTAGTACTGACGATATTTACCGAGATCTGATTTATAAC 6242  
QY 266 SerAlaProPheIleProMet---GlyValPheGlyHisIleAlaPro---ValLeuSer 283  
Db 6241 GGTACGACTTACACTCACTGATTAATGATCGATCCGATATCGAGCAATACCTGTTT 6182  
QY 284 IleGlySerLeuSerLysSerTrpIleValProGlyTyrAspGlyLeuGlyTrpValAlaVal 303  
Db 6181 ATTTCCGGGCTTCCAAATCATATCGATGACGGGTTCGCGGATTTGGTATCGCGCC--- 6125  
QY 304 TyrAspProThrLysIleLeuGlu-----LysThrLysIleSerThrSerIleThrAsn 321  
Db 6124 ---GGTCCGGAAGAGCTGATTCAGGCCCATGCGACCTTTTATTAGCCACACGACCTCTAAT 6068  
QY 322 TyrLeuAsnValSerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeu 341  
Db 6067 CCGGAGCAGATTTTCGAATACCCCGGTGGCAGCTTTAACTGGCGATCAGCAGGTGTT 6008  
QY 342 GluAsnThrLysAlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGlu 361  
Db 6007 GAAAAGATGCGCGCTGCTTTTGAAGAAGCGGTG-----AAT 5972  
QY 362 IleCysTyrArgGluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGly 381  
Db 5971 CTTTTCTATGATCTTCTGCGAGATATCCCGGTTTCGATATGGGAGATAAAACCGCAAGGC 5912  
QY 382 SerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspIle 401  
Db 5911 GCTTCTATCTTTTCCCGAATTAATTAAGCGTGGCGCTCAATTGAGTCATATGTTACGGTT 5852  
QY 402 ---AspPheCysCysLysLeuAlaLysGluLeuSerValIleLeuCysProGlySerVal 420  
Db 5851 GATGATTTTATCAGTCGACTGTGTGACCGAAACCGGGTTCGCAATGTTCTTGGACGGCG 5792  
QY 421 -LeuGlyMetGluAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAs 440  
Db 5791 TTTGGCATGCGCGATCATGCGCGGATAGTTATTGTTAAAGATTGGCCAGTCTGAAAGA 5732  
QY 440 pGlyLeuGluArgValLysSerPheCysGlnArgAsn 452  
Db 5731 GGCCCGCCGCGCTATCCGGAGCTTTGTTGTTAAATAAT 5695

## RESULT 12

US-08-599-171A-18  
; Sequence 18, Application US/08599171A  
; Patent No. 5814473  
; GENERAL INFORMATION:  
; APPLICANT: WARREN, Patrick V.  
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068

COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/599,171A  
; FILING DATE: Concurrently  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HERRON, CHARLES J.  
; REGISTRATION NUMBER: 28,019  
; REFERENCE/DOCKET NUMBER: 331400-38  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1122 NUCLEOTIDES  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: GENOMIC DNA  
; US-08-599-171A-18

Alignment Scores:  
Pred. No.: 1-07e-24 Length: 1122  
Score: 308.50 Matches: 103  
Percent Similarity: 44.82% Conservatives: 70  
Best Local Similarity: 26.68% Mismatches: 173  
Query Match: 12.91% Indels: 40  
DB: 12 Gaps: 12

US-10-019-783-1 (1-461) x US-08-599-171A-18 (1-1122)

QY 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspSerValPheProAlaPhe 94  
Db 51 TCAGGCCCAAGACTACGAGACGCTAGTACACATGAGATAGGAGACC-----CGATT 104  
QY 95 ArgThrAlaValGluAlaGlu-AspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAs 114  
Db 105 AGAACCGTCTCCCAAGGTAATGGAAGCTTGGAACGTGGGTGAAGAAAGACGCTTC-- 162  
QY 114 nCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSe 134  
Db 163 -TTCTACACCCCTGCTCTGGGACTCTGGGAACCTCAGGAAAGATATCGAGCTTTTACAG 221  
QY 134 rGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrG1 154  
Db 222 GAAAGAGTACAGCTTGAAGTTCTCCAGAGAGATCATCGTAACCTACCGGAACCTTCGG 281  
QY 154 nAlaLeuGluValLeuIleProValLeuAlaGlnThrAlaGlyAlaAsnLeuLeuPr 174  
Db 282 AGCGTTTCTCGTAGCTACCGCGTA--ACACTAAATGGCGAGAGAGATATCTCTCC 338  
QY 174 cArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHi 194  
Db 339 AGACCCCTCTTACCCCTGTTTCAAAAACCTTTCCTTACCTCTTACGCTTCAGCCGGTTT 398  
QY 194 sPheAspLeuIleProAspLysGlyTyrGluLeuAspLysLeuGluSerIleAl 214  
Db 399 CGTAAACGTTGACAGGAACGAATTACGAAGTAAGGAAGATGATAGAGACATTGA 458  
QY 214 aspLysAsnThrThrAlaMetValLeuIleAsnProAsnProCysGlySerValTy 234  
Db 459 TCGGAAA-----GCCCTTCACATTTTCCTCGCTCAAAACCTTACCGGCACACTCTA 509  
QY 234 rSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValI1 254  
Db 510 CTCACCTGAACCTGAGGAAGTTCGGAGTACTGCGAAGAGAGAGGTATGTACTTCAT 569  
QY 254 eAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVa 274  
Db 570 ATCCGACGAGATTACCGGACTCGTTTACGAGGTAGGAG----- 612  
QY 274 lPheGlyHisIleAlaProValLeuSer-----IleGlySerLeuSerLy 289  
Db 613 -----CACACAGACTTGAGTCTCTGACAGGCTATGTGCTAAACCGGTTTCTTAA 665  
QY 289 sSerTrpIleValProGlyTyrPArgLeuGlyTyrPValAlaValTyrAspProThrLysI1 309  
Db 666 GTACTTCTGTATGCCAGGTTTCAGGATAGGTGGATGATGTT-----CCGGAAGACT 719  
QY 309 eLeuGluLysThrLysIleSerThrIleThrAsnTyrLeuAsnValSerThrAspPr 329  
Db 720 CGTGAGAAAGCGGGAATAGTAATTTCAG-----AACGTATTATATATCTGC 764  
QY 329 oAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePh 349  
Db 765 CCGGACGCTCAGTCACTAGCGCC-----GCCCTTGGGCTTTGATTACGAGTATT 815  
QY 349 eLysArgIleLeuGlyLeuLysGluSerSerGluLeuCysTyrArgGluIleLysGl 369  
Db 816 GGAGAAGGTAGAAAAACCTTTGAAGAGAGAGGAACTTCCTTTATGCGGAACCTCAAA-- 873  
QY 369 uAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLe 389  
Db 874 -----AAACTCTCAAGATAGACCGGAACCTTCAGGAGCTTTTACGATCGGCAACAT 929  
QY 389 uAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLy 409  
Db 930 AAGT-----GATTACTCCACAGATAGCTACGAATTTGCTTTAAACCTTTTAAAG 977  
QY 409 sGluLysSerValIleLeuCysProGlySerValLeuGly-----MetGluAsnTr 426  
Db 978 GGAGCGGAGGTCGCGGTAAACCGCGGTGAGCTTTGGAATAAAACAAACGAAGAGTA 1037  
QY 426 pValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVally 446

Db 1038 TATAAGGTTTCTTATACGAAAAAGATAGAAGACTTAAAGGAGGCGCTTGAAGAGTAAA 1097  
QY 446 sSerPheCysGlnArg 451  
Db 1098 GAAGTCTTAGAGAAG 1113  
RESULT 13  
US-08-646-590B-18  
; Sequence 18, Application US/08646590B  
; Patent No. 5962283  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Patrick V.  
; APPLICANT: Swanson, Ronald V.  
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/646,590B  
; FILING DATE: 08-May-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/599,171  
; FILING DATE: 09-FEB-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/01094  
; FILING DATE: 21-January-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hallie, Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 09010/017001  
; TELEPHONE: 619/678-5070  
; TELEFAX: 619/678-5099  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1122 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 1...1119  
; US-08-646-590B-18  
Alignment Scores:  
Pred. No.: 1,07e-24 Length: 1122  
Score: 308.50 Matches: 103  
Percent Similarity: 44.82% Conservative: 70  
Best Local Similarity: 26.68% Mismatches: 173  
Query Match: 12.91% Indels: 40  
DB: 2 Gaps: 12  
US-10-019-783-1 (1-461) x US-08-646-590B-18 (1-1122)

QY 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspSerValPheProAlaPhe 94

Db 51 TCAGGCCCAAGACTACGAGACGCTAGTACACATGAGATAGGAGACC-----CGATT 104

QY 95 ArgThrAlaValGluAlaGlu-AspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAs 114

Db 105 AGAACCGTCTCCAGAGTAATGGAAGCTCTGGAAGCTGGAAGGAAAGACGTTTC-- 162  
Qy 114 nCystyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSe 134  
Db 163 -TTCTACACCCCTGCTCGGACTCTGGAACTCAGGAAGGATATCGGAGTTTACAG 221  
Qy 134 rGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrG1 154  
Db 222 GAAAAGTACAGCGTTGAAGTTTCTCCAGAGAGATCATGTAACCTACCGGAACCTCGG 281  
Qy 154 nAlaIleGluValIleLeuProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPr 174  
Db 282 AGCGTTTCTCGTAGCTACGCCGTA--ACACTAAATCGCGGAGAGAAATAATCTCTCCC 338  
Qy 174 oArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHi 194  
Db 339 AGACCCCTCTACCCCTCTTACAAAACCTTTGCCTACTCTTAGACGCTCAGCCGGTTT 398  
Qy 194 sPheAspLeuLeuProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAl 214  
Db 399 CGTAAACCTTGACAGGAAGAACGATACGAAGTAAGGAAGAGATGATAGAGACATGA 458  
Qy 214 aAspLysAsnThrThrAlaMetValIleLeuAsnProAsnAsnProCysGlySerValTy 234  
Db 459 TCGGAAA-----GCCCTTCACATTTCTCGCTCAAAACCCCTACGGGCACACTCTA 509  
Qy 234 rSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValI1 254  
Db 510 CTCACCTGAAACCTTGAGGAAGTTCGGAGTACTCGCAAGAGAGAGGTATGATCTTCAT 569  
Qy 254 eAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVa 274  
Db 570 ATCCGACGAGATTTACACGAGCTCGTTTACGAAGGTAGGGAG----- 612  
Qy 274 lPheGlyHisIleAlaProValLeuSer-----lleglySerLeuSerLy 289  
Db 613 -----CACAGACACTGAGTCTCTGACAGGCTATTGTCATAACGGGTTTCTAA 665  
Qy 289 sSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThrLysI1 309  
Db 666 GTACTTCTGTATGCAGGTTTCAGGATAGGGTGGATGATGTT-----CCGGAAGAACT 719  
Qy 309 eLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPr 329  
Db 720 CGTGAAAGGCGGAAATAGTAATTCAG-----AACGTATTATATCTGC 764  
Qy 329 oAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePh 349  
Db 765 CCCGACGCTCAGTCACTAGTACGCC-----GCCCTTGAGGCTTTTGATTCAGGATATT 815  
Qy 349 eLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysG1 369  
Db 816 GGAGAAGGTAGAAAACCTTTGAAGAGAGAGAGAACTTCCTTTATGGGGAAGTAA-- 873  
Qy 369 uAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLe 389  
Db 874 ----AAACTCTCAAGATAGACGCGAAACCTCAGGAGCTTTTACGTATGGGCAAAACAT 929  
Qy 389 uAsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLy 409  
Db 930 AAGT-----GATTACTCCACAGATAGCTACGAATTTGCTTTAAACACTTTTAAG 977  
Qy 409 sGluGluSerValIleLeuCysProGlySerValLeuGly-----MetGluAsnTr 426  
Db 978 GGAGCGAGGGTGGCGGTAAACCGCGGGGTGAGCTTTGGAAGAAACAAACAGGAGGTA 1037  
Qy 426 pValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVally 446  
Db 1038 TATAAGGTTTGTCTATACGAGAAGATAGAAAGAACTTAAGGCGGCGTTGAAGAGATAA 1097  
Qy 446 sSerPheCysGlnArg 451

Db 1098 GAAGTTCTTAGAGAAG 1113  
RESULT 14  
US-09-069-226-18  
Sequence 18, Application US/09069226  
Patent No. 6013509  
GENERAL INFORMATION:  
APPLICANT: WARREN, Patrick V.  
TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILPILLAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/069,226  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/599,171  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: HERRON, CHARLES J.  
REGISTRATION NUMBER: 28,019  
REFERENCE/DOCKET NUMBER: 331400-38  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1122 NUCLEOTIDES  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: GENOMIC DNA  
US-09-069-226-18  
Alignment Scores:  
Pred. No.: 1,07e-24 Length: 1122  
Score: 308.50 Matches: 103  
Percent Similarity: 44.82% Conservative: 70  
Best Local Similarity: 26.68% Mismatches: 173  
Query Match: 12.91% Indels: 40  
DB: 12 Gaps: 3  
US-10-019-783-1 (1-461) x US-09-069-226-18 (1-1122)  
Qy 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 94  
Db 51 TCAGGCCCAAGAGTACCAAGAGCTAGTACATGCGATAGAGAGGCC-----CGATT 104  
Qy 95 ArgThrAlaValGluAlaGlu-AspAlaValAlaAlaLeuAspThrGlyGlnPheAs 114  
Db 105 AGAACCGTCTCCAGAGTAATGGAAGCTCTGGAAGCTGGAAGGAAAGACGTTTC-- 162  
Qy 114 nCystyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSe 134  
Db 163 -TTCTACACCCCTGCTCGGACTCTGGAACTCAGGAAGGATATCGGAGTTTACAG 221  
Qy 134 rGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrG1 154  
Db 222 GAAAAGTACAGCGTTGAAGTTTCTCCAGAGAGATCATGTAACCTACCGGAACCTCGG 281  
Qy 154 nAlaIleGluValIleLeuProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPr 174

```

Db 282 AGCGTTCTCGTAGCCTACCGCGTA---ACACTAAATGCGGAGAGAGATAATCTCTCC 338
QY 174 oArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHi 194
Db 339 AGACCCCTCTTACCTCTGTTACAAAACCTTGCCTACCTCTTAGAGCTCAGCGGTTT 398
QY 194 sPheAspLeuLeuProAspLysGlyTrpGluLeuAspLeuAspSerLeuGluSerIleAl 214
Db 399 CGTAAACCTTGAACAAGAACGAATACGAAGTAAGAAAGAGATGATAGAACATTTGA 458
QY 214 aaSPlysAsnThrAlaMetValIleIleAsnProAsnProCysGlySerValTy 234
Db 459 TCGGAAA-----GCCCTTCACATTTCCCTCGCCTCAAAACCTTACGGGCACACTCTA 509
QY 234 rSerTyrAspHisLeuAlaLysValaGluValaAlaArgLysLeuGlyIleLeuValIl 254
Db 510 CTCACCTGAACCCCTGAAGAACTTCGGAGTACTCGGAAGAGAGAGGTATGTACTTCAT 569
QY 254 eAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVa 274
Db 570 ATCCGAGAGATTTACCGGACTCGTTTACGAAGTAGGAG----- 612
QY 274 lPheGlyHisIleAlaProValLeuSer-----IleGlySerLeuSerLy 289
Db 613 -----CACACAGCACTTGAGTCTCTGCACAGGCTATTGTCATAAACGGGTTTCTAA 665
QY 289 sSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspProThrLysIl 309
Db 666 GTACTTCTGTATGCCAGGTTTCAGATAGGTTGATAGT-----CCGGAAGAACT 719
QY 309 eLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPr 329
Db 720 CGTGAGAAAGCGGAATAGTAATCAG-----AACGTATTATATCTGC 764
QY 329 oAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePh 349
Db 765 CCCGACGCTCAGTCAGTACGCC-----CCCTTCAGGCTTTTGATTACGAGTATT 815
QY 349 eLysArgIleIleGlyLeuLysGluSerSerGluIleCysTyrArgGluIleLysGl 369
Db 816 GGAGAGGTAAAGAAACCTTTGAGAGAGAGAGGAACTTCTTTATGGGAACGAA-- 873
QY 369 uAsnLysTyrIleThrCysProHisLysProGlySerMetPheValMetValLysLe 389
Db 874 ----AAACTCTTCAAGATAGACGCGAAACCTCAGGAGCTTTTACGTATGGCAACAT 929
QY 389 uAsnLeuHisLeuLeuGluGluIleHisAspAspLeuAspPheCysCysLysLeuAlaLy 409
Db 930 AAGT-----GATTCTCCACAGATAGCTACGAATTTGCTTTTAAACATTTTAAAG 977
QY 409 sGluGluSerValIleLeuCysProGlySerValLeuGly-----MetGluAsnTr 426
Db 978 GGAGCGGAGGTGGCGGTAAACGCCCGGGTGGACTTTGGAAAAACAAACAGAGGAGTA 1037
QY 426 pValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLy 446
Db 1038 TATAAGGTTTGTATTATACGAGAAAGATAGAAAGAACTTAAAGGCGCGTTGAAGGATAA 1097
QY 446 sSerPheCysGlnArg 451
Db 1098 GAAGTTCTTAGAAG 1113

```

## RESULT 15

```

US-09-412-184-18
; Sequence 18, Application US/09412184
; Patent No. 6268188
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/412.184
; FILING DATE:
; CLASSIFICATION:

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646.590
; FILING DATE: 08-May-1996
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Hallie, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099

```

```

; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1122 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1119
; US-09-412-184-18

```

```

Alignment Scores:
Pred No.: 1,07e-24 Length: 1122
Score: 308.50 Matches: 103
Percent Similarity: 44.82% Conservative: 70
Best Local Similarity: 26.68% Mismatches: 173
Query Match: 12.91% Indels: 40
DB: 3 Gaps: 12

```

US-10-019-783-1 (1-461) x US-09-412-184-18 (1-1122)

```

QY 75 SerGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 94
Db 51 TCAGGCCCAAGTACGAACACGTAGTACACATGGAGTAGAGAGACC-----CGATTT 104
QY 95 ArgThrAlaValGluAlaGlu-AspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAs 114
Db 105 AGAACCGTCTCCCAAGGTATGGAGACTCTGGAACTGCGGTGAGGAAAGAACGTTCC-- 162
QY 114 nCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSe 134
Db 163 -TTCTACACCCCTGCTCTGGAGACTCTGGAACTCAGGAAAGGATATCGGAGTTTACAG 221
QY 134 rGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGl 154
Db 222 GAAAAAGTACAGCGTTTCAAGTTTCTCCAGAGAGAGTCATCGTAACCTACCGGAACITCGG 281
QY 154 nAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPr 174
Db 282 AGCGTTTCTCGTAGCCTACCGCGTA---ACACTAAATGCGGAGAGAGATATCTCTCC 338

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 23:39:05 ; Search time 408.614 Seconds  
(without alignments)  
3600.055 Million cell updates/sec

Title: US-10-019-783-1

Perfect score: 2390

Sequence: 1 MVHQNGHGAAAAAANKS.....LERYKSFQCKKKKNSINGC 461

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2141354 seqs, 1595478879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10019783/runat\_07112003\_161647\_19888/app\_query.fasta\_1.1358  
-DB=Published Applications NA -QFMT=fastap -SUFFX=p2n.rnpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Dlosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10019783 @CGN 1 1 541 @runat\_07112003\_161647\_19888  
-NCPV=6 -ICPV=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100\_19988  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/pct\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09D\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1124	47.0	1389	10	US-09-938-842A-1984

	2	1023.5	42.8	1338	10	US-09-938-842A-1869	Sequence 1869, Ap
C	3	649	27.2	858	9	US-09-770-445-608	Sequence 608, App
C	4	644	26.9	2712	11	US-09-919-039-379	Sequence 379, App
	5	644	26.9	2754	10	US-09-880-107-3720	Sequence 3720, Ap
	6	644	26.9	2921	14	US-10-198-846-11998	Sequence 11998, A
	7	611	25.6	1631	14	US-10-198-846-11931	Sequence 11931, A
C	8	390	16.3	870	14	US-10-198-846-8896	Sequence 8896, Ap
	9	361.5	15.1	1185	10	US-09-905-173-39	Sequence 39, Appl
C	10	355.5	14.9	1224	14	US-10-156-761-4891	Sequence 4891, Ap
	11	355.5	14.9	9025608	14	US-10-156-761-4891	Sequence 1, Appl
	12	355	14.9	3309400	10	US-09-738-626-1	Sequence 1, Appl
	13	354	14.8	1257	10	US-09-738-626-3134	Sequence 3134, Ap
	14	331	13.8	370	10	US-09-878-574-2935	Sequence 2935, Ap
	15	323.5	13.5	9025608	14	US-10-156-761-1	Sequence 1, Appl
	16	308.5	12.9	1122	10	US-09-905-173-18	Sequence 18, Appl
	17	308.5	12.9	1122	14	US-10-060-432-18	Sequence 18, Appl
	18	307	12.8	1206	14	US-10-156-761-1996	Sequence 1996, Ap
	19	306	12.8	1209	14	US-10-156-761-3272	Sequence 3272, Ap
	20	304.5	12.7	1409	10	US-09-971-536-11	Sequence 11, Appl
	21	301	12.6	1830121	14	US-10-329-960-1	Sequence 1, Appl
	22	300	12.6	1247	10	US-09-971-536-12	Sequence 12, Appl
	23	299	12.5	1194	14	US-10-156-761-4501	Sequence 4501, Ap
C	24	299	12.5	21252	10	US-09-070-927A-94	Sequence 94, Appl
	25	297.5	12.4	837	14	US-10-198-846-12727	Sequence 12727, A
	26	295.5	12.4	1170	12	US-10-288-930-5	Sequence 5, Appl
	27	294	12.3	1170	9	US-09-967-645-2	Sequence 2, Appl
C	28	293	12.3	3614	10	US-09-070-927A-185	Sequence 185, App
	29	289	12.1	1209	10	US-09-738-626-911	Sequence 911, App
	30	281.5	11.8	1107	12	US-10-288-930-4	Sequence 4, Appl
	31	265.5	11.1	1197	10	US-09-905-173-21	Sequence 21, Appl
	32	265.5	11.1	1197	14	US-10-060-432-21	Sequence 21, Appl
	33	261.5	10.9	1221	10	US-09-938-842A-1197	Sequence 1197, Ap
C	34	245	10.3	446	9	US-09-770-444-786	Sequence 786, App
	35	243	10.2	1975	14	US-10-284-985-4	Sequence 4, Appl
	36	243	10.2	1975	14	US-10-284-014-1	Sequence 1, Appl
	37	242.5	10.1	1076	10	US-09-070-927A-835	Sequence 835, App
	38	242	10.1	730	9	US-09-770-149-110	Sequence 110, App
	39	240.5	10.1	13884	10	US-09-070-927A-341	Sequence 341, App
	40	239	10.0	458	9	US-09-770-444-496	Sequence 496, App
C	41	239	10.0	870	14	US-10-198-846-2949	Sequence 2949, Ap
	42	236	9.9	1572	10	US-09-908-928-3	Sequence 3, Appl
	43	236	9.9	2127	10	US-09-908-928-1	Sequence 1, Appl
	44	228	9.5	764	10	US-09-974-300-292	Sequence 292, App
	45	222.5	9.3	1101	14	US-10-156-761-3147	Sequence 3147, Ap

#### ALIGNMENTS

RESULT 1  
US-09-938-842A-1984  
Sequence 1984, Application US/09938842A  
Patent No. US20020160378A1

GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Zhu, Xun  
APPLICANT: Wang, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SCRIPL300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 1984  
LENGTH: 1389  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana



Db 55 -----AAAGAAACGGTCCACGAGTGTGCGCGCGGTGACGTTGAGAAAGCTTCT 108  
QY 66 TyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeuProLeuAlaHis 85  
Db 109 TTTGGATGTTTAAAAACATGCACCATGAACAGTGAAGAACCATTTTGTTCACACTCCC 168  
QY 86 GlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaVal 105  
Db 169 GCGAGCCCTCGCCCACTCCACTTCAGGATTCGCCGGAAGCCGAGAGCCGTTGCC 228  
QY 106 AlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAla 125  
Db 229 GCAGCTGCACGCTCGCGATGCTTAACCTTACGCCACCCAGCCCTGGAGTTTCAAGCT 289  
QY 126 ArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAsp 145  
Db 289 AAGAGCGCGTGGCTGAATTTTAAACGGAGAACTTCCGACGAAGCTGAAGCCGAGGAT 348  
QY 146 ValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGln 165  
Db 349 GTGTATATCACCGAGGAGTGAACCAAGCCATAGAGATCGTATAGATTCCTTCCCGGA 408  
QY 166 ThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArg 185  
Db 409 AATCCATCCGCCCAACATTTACTTCCAGCGCGGATATCTCTCACTACGATGCTCGTCT 468  
QY 186 AlaPheAsnLysLeuGluValAlaArgHisPheAspLeuIleProAspLysGlyTrpGluIle 205  
Db 469 GTCTATAGCGGCTCGAGATTCGCAATACGATCTTCTCCCGAGAGTGATGGAAATC 528  
QY 206 AspIleAspSerLeuGluSerIleAlaAspLysAsnThrAlaMetValIleIleAsn 225  
Db 529 AATCTCGATGGCTCGAGCGCGCTCGGATGAGATACCGTCGCAATGGTAAATCATCAAC 588  
QY 226 ProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluVal 245  
Db 589 CCCAACATCCATGTGGAAGCTTACACCTACGACCATCTCAACAGAGTCGCGGAGATG 648  
QY 246 AlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGly 265  
Db 649 GCTAGAAACCTGGTAT 708  
QY 266 SerAlaProPheIleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGly 285  
Db 709 GACAGCCCTTTATTCCTCCAGGAGTTCGATCAATAGCTCCGCTGATCAGCTCGGA 768  
QY 286 SerLeuSerLysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAsp 305  
Db 769 TCCATATCCAAAGGATGGTCAACCCAGCTGGAGATTTGGCTGATCGCCATCGAAGCAT 828  
QY 306 ProThrLysIleLeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 325  
Db 829 CCTATGTATCTTTGTATCTACAGGGGTAGTTCAGGCAATAGAGGATTTCTTGTATTA 888  
QY 326 SerThrAspProAlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLys 345  
Db 889 ACTCCAGCCCTTATTTATCTCCAGGAGCATCTTCTGTATATATATATATATATATAT 948  
QY 346 AlaAspPhePheLysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArg 365  
Db 949 AAAGAGTTCTTCGAGAAGCAAGATCAACCCATGAGACCGCAAGCTTCATGTGAG 1008  
QY 366 GluIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 385  
Db 1009 AGGCTCAAGGATATTCCTTGTCTCTTTGTCCCAAGAAACCCGATCTTGTCTCTATTA 1068  
QY 386 MetValLysLeuAsnLeuHisLeuLeuGluIleHisAspAspIleAspPheCysCys 405  
Db 1069 TGGTTGAGCTTGACACATCATGTTGATATATATCAAAATGATTTTGAATTTCTGACG 1128  
QY 406 LysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsn 425

Db 1129 AAGCTAGTTAGTAGGAGAGTCTTATCTTATATACGAGAGTGGCTCTAGGGGAGAGAA 1188  
QY 426 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgVal 445  
Db 1189 TGGGTGAGGATATCGATAGGAACCGACGAATCAGTGTGACAGAAATATTTGACAGACTA 1248  
QY 446 LysSerPheCysGlnArgAsn 452  
Db 1249 AAAGTTTCTATGATCGTCAT 1269  
RESULT 3  
US-09-770-445-608/c  
; Sequence 608. Application US/09770445  
; Patent No. US2002023281A1  
; GENERAL INFORMATION:  
; APPLICANT: Goriach, Jörn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kricker, Maja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE REFERENCE: 2023US (PARA-012PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,445  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,472  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 608  
; LENGTH: 858  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-445-608  
Alignment Scores:  
Pred. No.: 3.8e-65 Length: 858  
Score: 649.00 Matches: 118  
Percent Similarity: 70.67% Conservative: 41  
Best Local Similarity: 52.44% Mismatches: 66  
Query Match: 27.15% Indels: 0  
Gaps: 0  
US-10-019-783-1 (1-461) x US-09-770-445-608 (1-858)  
QY 230 CysGlySerValTyrSerTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeu 249  
Db 858 TGTGGAGATGTTTCTCTCCCAACATCTTCAAAAGATTCAGAGACGGCTTGCAAGCTT 799  
QY 250 GlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPhe 269  
Db 798 GGTATACTTGTGATCGACAGCAAGTCTATGACCATTTTTCCTTTGGGATAAACCATTT 739  
QY 270 IleProMetGlyValPheGlyHisIleAlaProValLeuSerIleGlySerLeuSerLys 289  
Db 738 GTGTCATGGCAGAGTTTCAGAGCTAGTGCCTGTATTTGTTAGTCTATATCTATAA 679  
QY 290 SerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProThrLysIle 309

Db 678 AGATGGTTTCTCGTGAATGAGACTGGTTGGATGGTGAATCTTGACCTCATGGCATC 619  
Qy 310 LeuGluLysThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPro 329  
Db 618 ATGAAGATTCTGGGTTTGTTCAGACTCTTATCAATGTGTCAACATGTGCGCGATCCT 559  
Qy 330 AlaThrPheValGlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePhe 349  
Db 558 GCAAGTTTATTCAGGGAGCAATGCTGATATCATTTGGGATACAAAGGAGAGTTTCTTC 499  
Qy 350 LysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGlu 369  
Db 498 TCATCAAACTTGAAATGGTGAATAATGTGCAGAGATTTGTTATGAGGAGCTTATCAAG 439  
Qy 370 AsnLysTyrIleThrCysProHisLysProGluLysSerMetPheValMetValLysLeu 389  
Db 438 ATTCTTGCATCCTACCTACCAACCCGAGGGCTCAATGTTTCCAGCATGGTGAAGTTA 379  
Qy 390 AsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLys 409  
Db 378 AACTTTTCATCTACGAGATATCAGTATGATTTGGACTTCTGCTCCAGCTGGCTAAA 319  
Qy 410 GluGluSerValIleLeuCysProGlySerValLeuGluMetGluAsnTrpValArgIle 429  
Db 318 GAGGAATCTATGATCATCTACCAAGCTGTCAGCTGTTGGGCTTAAGAACTGGCTACGTATC 259  
Qy 430 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCys 449  
Db 258 ACCTTTGCAATGAGCTTGAGCTTCTATAGAGGGTTTCCAGGCTTAAGAACTTTACT 199  
Qy 450 GlnArgAsnLysLys 454  
Db 198 GAGAGACACTCCAG 184

## RESULT 4

US-09-919-039-379/c  
; Sequence 379, Application US/09919039  
; Publication No. US20030108871A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919,039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/222,113  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 379  
; LENGTH: 2712  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Inocyte ID No. US20030108871A1 335648.1C  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 236-590  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-919-039-379

Alignment Scores:  
Pred. No.: 9,098-64 Length: 2712  
Score: 644.00 Matches: 144  
Percent Similarity: 53.08% Conservative: 89  
Best Local Similarity: 32.80% Mismatches: 176  
Query Match: 26.95% Indels: 30  
DB: 11 Gaps: 10

US-10-019-783-1 (1-461) x US-09-919-039-379 (1-2712)

Qy 22 GlyHisAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGlu 41

Db 2544 GGTGGGAGAGCTCTGTGCCGGGAAAAATGAAAGGCAGAAAGGCC-----AGG 2497  
Qy 42 TrpAsnPheAlaArgGlyAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIle 61  
Db 2496 TGGTCT--GTGAGGCCCTCAGACATGCCAAGAAACTTTT-----AACCCCATC 2449  
Qy 62 ArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeu 81  
Db 2448 CAGGCAIT-----GTGGACAACATGAAGTGAACCAATCCAAACAAACCATGATT 2395  
Qy 82 ProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGlu 101  
Db 2394 TCCCTGTCCATTGGGAGCCCTACTGTGTGTGAAACCTGCCTACAGACCCCTGAAGTTACC 2335  
Qy 102 AspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGly 121  
Db 2334 CAGGCAATGAAGATGCCCTGGACTCGGCGAAATATATATGCTATGCCCATCCATCGGC 2275  
Qy 122 LeuProAlaAlaArgSerAlaValAlaGlu-----HisLeuSerGlnGlyValProTyr 139  
Db 2274 TTCTATCCAGTCGGGAGGAGATTGCTTCTTATTACCACTGTCTCTAGGCCA----- 2224  
Qy 140 LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159  
Db 2223 CCCCTAGAACTAAGACGCTATTCTGACAAGTGGCTGCAGCAAGCTATTGACCTTGT 2164  
Qy 160 IleProValAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 179  
Db 2163 TTAGCTGTGTGGCC--AACCCAGGGCAAAACATCCTGGTTCCTCAAGACCTGTTCTCT 2107  
Qy 180 AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199  
Db 2106 CTCTACAGACTCTGCTGAGTCTATGGAATTTAGGTCAAACTCTACAATTTGTGGCA 2047  
Qy 200 AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219  
Db 2046 GAGAAATCTTGGGAAATTTGACTGAAACACTGGATATCTAATTTGATGAAAAGACAGT 1987  
Qy 220 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239  
Db 1986 TGTCTCATTTGTCATCAATCAATCAAAACCCCTGGGTCAGTGTTCAGAAACGTCATCTT 1927  
Qy 240 AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyr 259  
Db 1926 CAGAAGATTCTGGCAGTGGCTGCACGGCAGTGTGTCCCACTCTTACGTATGAGATCTAT 1867  
Qy 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279  
Db 1866 GGAGACATGGTGTGTTTCGGATTGCAATATGAACCACTGGCCCACTCAGCACCGATGTC 1807  
Qy 280 ProValLeuSerIleGlySerLeuSerLysSerTyrIleValProGlyTyrArgLeuGly 299  
Db 1806 CCATCTCTCTGTGGAGGGCTGGCAAGCCCTGGCTGTCTCTGGCTGGAGGTTGGGC 1747  
Qy 300 TrpValAlaValTyrAspProThrLysIleLeuGluLys-----Thr 313  
Db 1746 TGGATCCTCATTCATGACCGAAGAGACATTTTGGCAATGAGATCCGAGATGGCTGGTG 1687  
Qy 314 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheVal 333  
Db 1686 AAGCTGAGTCAGCGCATTTTGGGA-----CCCTGTACCATTTGTC 1648  
Qy 334 GlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPheLysArgIleIle 353  
Db 1647 CAGGAGCTCTGAAGACATCTTATGTCACCCCGGAGAGCTTTTACCACAACACTCTG 1588  
Qy 354 GlyLeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIle 373  
Db 1587 AGCTTCTCAAGTCCAATGCTGATCTCTGTGTGGGGGCTGCTGCCATCCCTCGACTC 1528  
Qy 374 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 393  
Db 1527 CGG---CCAGTCGGCCCTTCTGGGGCTATGCTACCTCATGTTGGAATTCAGATGGACAT 1471

QY 394 LeuGluGluIleHisAspPheCysLysLeuAlaLysGluGluSerVal 413  
 Db 1470 TTCCAGAAATTCAGAACAGATGGAGTTACGGAGCGTTAGTGTGAGCGTCTGTC 1411  
 QY 414 IleLeuCysProGlySerValLeuGluMetGluAsnTrpValArgIleThrPheAlaCys 433  
 Db 1410 CACTGCTCCACGACACGCTGTTGAGTACCGCAATTTTCATCCGAGTGGTCAACAGTC 1351  
 QY 434 ValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452  
 Db 1350 CCCGAGGTGATGCTGGAGCGTGCAGCCGATCCAGGAGTTCTGTGAGCAGCAC 1294

## RESULT 5

US-09-880-107-3720  
 ; Sequence 3720, Application US/09880107  
 ; Patent No. US20020142981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horne, Darci T.  
 ; APPLICANT: Vockley, Joseph G.  
 ; APPLICANT: Scherif, Uwe  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 ; FILE REFERENCE: 44921-5028-WO  
 ; CURRENT APPLICATION NUMBER: US/09/880,107  
 ; PRIOR FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/211,379  
 ; PRIOR FILING DATE: 2000-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/237,054  
 ; PRIOR FILING DATE: 2000-10-02  
 ; NUMBER OF SEQ ID NOS: 3950  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 3720  
 ; LENGTH: 2754  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X52520  
 US-09-880-107-3720

## Alignment Scores:

Pred. No.: 9,326-64 Length: 2754  
 Score: 644.00 Matches: 144  
 Percent Similarity: 53.08% Conservative: 89  
 Best Local Similarity: 32.80% Mismatches: 176  
 Query Match: 26.95% Indels: 30  
 Db: 10 Gaps: 10

US-10-019-783-1 (1-461) x US-09-880-107-3720 (1-2754)

QY 22 GlyHisAlaAlaAlaAsnGlyLysSerAsnGlyHisAlaAlaAlaAlaValGlu 41  
 Db 169 GTGGGAGAGCTCTGTCCGGGAAATATGAAGCAGAAAGGCC-----AGG 216  
 QY 42 TrpAsnPheAlaArgGlyLysAspGlyIleLeuAlaThrGlyAlaLysAsnSerIle 61  
 Db 217 TGGTCT---GTGAGGCCCTCAGACATGCCCAAGAAACTTTC-----AACCCCATC 264  
 QY 62 ArgAlaIleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeu 81  
 Db 265 CGAGCCATT-----GTGGCAACATGAAGTGAACCAAAATCCAAACCAACCATGATT 318  
 QY 82 ProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGlu 101  
 Db 319 TCCCTGTCCATTGGGACCTTACTGTGTTTGGAAACCTGCCTACAGACCTTGAAGTTACC 378  
 QY 102 AspAlaValAlaAlaAlaLeuArgThrGlyGlnPheAsnCysTyrAlaAlaGlyValGly 121  
 Db 379 CAGCAATGAAGATGCCCTGACTCGGCGAAATATATGCTATGCTATGCCCATCCATCCGC 438  
 QY 122 LeuProAlaAlaArgSerAlaValAlaGlu-----HisLeuSerGlnGlyValProTyr 139  
 Db 439 TTCCTATCCAGTGGGAGAGATGCTTCTTATTACCACTGCTCTGAGGCA-----489

QY 140 LysLeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159  
 Db 490 CCCCTAGAGCTAGGACGTCATTCTCACAAGTGGCTGCAGCAAGTATTGACCTTGT 549  
 QY 160 IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 179  
 Db 550 TTACTCTGTGTGGC---AACCCAGGGCAGAAATCTCTGTTCCAGACCTGTTCTCT 606  
 QY 180 AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199  
 Db 607 CTCTACAAGACTCTGGCTGAGTCTATGGGAATTGAGGTCAAATCTACAAATTTGTGCCA 666  
 QY 200 AspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219  
 Db 667 GAGAAATCTTGGGAATTGACCTGAAACAACTGAATATCTAATTGATGAAGAACAGACT 726  
 QY 220 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239  
 Db 727 TGTCTCATTTCAATATCCATCAAAACCCCTGTGGGTGAGTCTTACGAAACGTCATCT 786  
 QY 240 AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyr 259  
 Db 787 CAGAAGATTCTGGCAGTGGCTGCAGCGACTGTGCCCATCTTAGCTGATGAGATCTAT 846  
 QY 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279  
 Db 847 GGAGACATGCTGTTTTCGGATTGCAATATGAACCACTGCCCACTCAGCACCGATGTC 906  
 QY 280 ProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGly 299  
 Db 907 CCCATCTGCTCTGTGGAGGGCTGGCAAGCGCTGGCTGGTCTCTGGTGGAGTTGGGC 966  
 QY 300 TrpValAlaValTyrAspProThrLysIleLeuGluLys-----Thr 313  
 Db 967 TGGATCCTCATTCATGACCAAGACATTTTGGCAATGAGATCCGAGATGGCTGGTG 1026  
 QY 314 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProLathrPheVal 333  
 Db 1027 AAGCTGAGTCAGCGCATTTTGGCA-----CCCTGTACCATTTGTC 1065  
 QY 334 GlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIle 353  
 Db 1066 CAGGAGCTCTGAAAAGCATCTTATGTCGACCCCGGAGAGTTTACCACAACTCTG 1125  
 QY 354 GlyLeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIle 373  
 Db 1126 AGCTTCTCAAGTCCAATGCTGATCTCTGTATGGGGCTGGCTGCCATCCCTGGACTC 1185  
 QY 374 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 393  
 Db 1186 CGG---CCAGTCCGCCCTTCTGGGGCTATGTACCTCATGTTGGAATTGAGATGGAACAT 1242  
 QY 394 LeuGluGluIleHisAspIleAspPheCysLysLeuAlaLysGluGluSerVal 413  
 Db 1243 TTCCAGAAATTGAGAACGATGGAGTTCCCGAGCGGTAGTCTCGAGCAGTCTGTC 1302  
 QY 414 IleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCys 433  
 Db 1303 CACTGCTCCAGCAACGCTGTTTGGATACCCGAAATTTTCATCCGAGTGTGTCATCAGATC 1362  
 QY 434 ValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452  
 Db 1363 CCCGAGGTGATGATGCTGGAGCGCTGCAGCGGATCCAGAGTTCTGTGAGCAGCAC 1419

## RESULT 6

US-10-198-846-11998  
 ; Sequence 11998, Application US/10198846  
 ; Publication No. US2003009997A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lillie, James  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
 ; FILE REFERENCE: MRI-049  
 ; CURRENT APPLICATION NUMBER: US/10/198,846  
 ; CURRENT FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/306,220  
 ; PRIOR FILING DATE: 2001-07-18  
 ; NUMBER OF SEQ ID NOS: 14084  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11998  
 ; LENGTH: 2921  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-198-846-11998

Alignment Scores:  
 Pred. No.: 1,02e-63 Length: 2921  
 Score: 644.00 Matches: 144  
 Percent Similarity: 53.08% Conservative: 89  
 Best Local Similarity: 32.80% Mismatches: 176  
 Query Match: 26.95% Indels: 30  
 DB: 14 Gaps: 10

US-10-019-783-1 (1-461) x US-10-198-846-11998 (1-2921)  
 QY 22 GlyHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaValGlu 41  
 DB 199 GTGGGAGAGCTCTGTCGGGGAATAATGAAGCAGCAAGAGGCC-----AGG 246  
 QY 42 TTPAsnPhaAlaArgGlyLysAspGlyLeuAlaThrGlyAlaLysAsnSerIle 61  
 DB 247 TGTCTCT--GTGAGCCCTCAGACATGCGCCAGAAACCTTC-----AACCCATC 294  
 QY 62 ArgAlaIleArgTyLysLysSerAlaSerValGluGluSerGlyProArgProValLeu 81  
 DB 295 CGAGCCATT-----GTGGACAACATGAAGTGAACCAATCCAAACCAACCATGATT 348  
 QY 82 ProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGlu 101  
 DB 349 TCCTGTCTCATTGGGACCTACTGTTTGGAAACCTGCCTACAGACCTGAGGTACC 408  
 QY 102 AspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaGlyValGly 121  
 DB 409 CAGGCAATGAAAGATGCTGCTGACTCGGCAATATAATGCTATGCTCCCACTCCATCGGC 468  
 QY 122 LeuProAlaAlaArgSerAlaValAlaGlu-----HisLeuSerGlnGlyValProTy 139  
 DB 469 TTCCTATCCAGTCGGGAGGAGATGCTCTTATACCAGTCTGCTGAGGCA-----519  
 QY 140 LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValle 159  
 DB 520 CCCCTAGAGCTAAGGACGCTATTCTGACAGTGGCTGCAGCAAGCTATTGACCTTGT 579  
 QY 160 IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyPro 179  
 DB 580 TTAGTGTGTGGCC---AACCCAGGCAACATCTCTGGTTCAGACCTGGTTCTCT 636  
 QY 180 AsnTyGluAlaArgAlaAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199  
 DB 637 CTCCTACAGACTCTGGCTGAGTCTATGGGAATGAGTCAAACTCTACAATTGTGGCA 696  
 QY 200 AspLysGlyTyGluIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219  
 DB 697 GAGAAATCTTGGGAATTGACCTGAACAACTGGAATATCTAATTGTCNAAGACGCT 756  
 QY 220 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTySerTyAspHisLeu 239  
 DB 757 TGTCTCATTGTCAATATCCATCAACCCCTGTGGTGTGAGTGTTCAGCAACGTCATCT 816  
 QY 240 AlaLysValAlaGluValAlaArgTysLeuGlyIleLeuValIleAlaAspGluValTy 259

DB 817 CAGAGATTCTGGCAGTGGCTGCACGGCAGTGTGTCCCATCTTAGTGTGATGAGATCTAT 876  
 QY 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279  
 DB 877 GGAGACATGCTGTTTTTCGGAATGCAAAATATGAACACTGGCCACCTCAGCACCGATGTC 936  
 QY 280 ProValLeuSerIleGlySerLeuSerLysSerTPIleValProGlyTyPArgLeuGly 299  
 DB 937 CCATCTCTCTCTGGAGGGTGGCCAGCGCTGGCTGGTTCCTGGCTGGAGGTGGGC 996  
 QY 300 TTPValAlaValTyAspProThrLysIleLeuGluLys-----Thr 313  
 DB 997 TGGATCTCTCATTCATCCAGCAAGAGACATTTTTTGGCAATGAGATCCGAGATGGCGTGTG 1056  
 QY 314 LysIleSerThrSerIleThrAsnTyLysLeuAsnValSerThrAspProAlaThrPheVal 333  
 DB 1057 AAGCTGAGTCAGGCAATTTTGGCA-----CCCTGTACCATTTGTC 1095  
 QY 334 GlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIle 353  
 DB 1096 CAGGAGCTCTGAAAAGCATCTATGTGCGACCCCGGAGAGCTTTTACCACAACTCTG 1155  
 QY 354 GlyLeuLeuLysGluSerSerGluIleCysTyArgGluIleLysGluAsnLysTyIle 373  
 DB 1156 AGTTCCTCAAGTCCCAATGCTATCTGTATTGGGGCTGGCTGCCATCCCTGGACTC 1215  
 QY 374 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 393  
 DB 1216 CGG---CCAGTCCGCCCTCTCGGCTATGTACTCTCATGTTGGAATGAGATGCAACAT 1272  
 QY 394 LeuGluGluIleHisAspAspIleAspPheCysLysLysLeuAlaLysGluSerVal 413  
 DB 1273 TTCCCAAGATTTGAGAACGATGTGAGTTTCACGGAGCGGTTAGTTGCTGAGCAGTCTGTC 1332  
 QY 414 IleLeuCysProGlySerValLeuGlyMetGluAsnTyPValArgIleThrPheAlaCys 433  
 DB 1333 CACTGGCTCCAGCAACGCTGTTGAGTACCCGAATTCATCCGAGTGGTCACTACAGTC 1392  
 QY 434 ValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452  
 DB 1393 CCGCAGGTGATGATGTGGAGGCGTGCAGCCGAGTCCAGGAGTTCTGTGAGCAGCAC 1449

RESULT 7  
 US-10-198-846-11531  
 ; Sequence 11531, Application US/10/98846  
 ; Publication No. US20030099974A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lillie, James  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Wang, Youzhen  
 ; APPLICANT: Steinmann, Kathleen  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
 ; FILE REFERENCE: MRI-049  
 ; CURRENT APPLICATION NUMBER: US/10/198,846  
 ; CURRENT FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/306,220  
 ; PRIOR FILING DATE: 2001-07-18  
 ; NUMBER OF SEQ ID NOS: 14084  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11531  
 ; LENGTH: 1631  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1571  
 ; OTHER INFORMATION: n = A,T,C or G  
 ; US-10-198-846-11531

Alignment Scores:  
 Pred. No.: 2.7e-60 Length: 1631

Score: 611.00 Matches: 141  
 Percent Similarity: 51.71% Conservative: 86  
 Best Local Similarity: 32.12% Mismatches: 164  
 Query Match: 25.56% Indels: 48  
 DB: 14 Gaps: 11

US-10-019-783-1 (1-461) x US-10-198-846-11531 (1-1631)

QY 22 GlyHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaValGlu 41  
 DB 181 GTGGGAGAGCTCTCTGCGGGAATAATCAAGGCGAAGAGGCC-----AGG 228  
 QY 42 TrpAsnPheAlaArgGlyLysAspGlyIleLeuAlaThrThrGlyAlaLysAsnSerIle 61  
 DB 229 TGGTCT--GTGAGGCCCTCAGACATGGCCACAAACTTTC-----AACCCATC 276  
 QY 62 ArgAlaIleArgTyrLysIleSerAlaSerValGluLysSerGlyProArgProValLeu 81  
 DB 277 CGAGGCCAAT-----GTGGACACATCAAGGTGAACCAATCAACAAACCAATGATT 330  
 QY 82 ProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGlu 101  
 DB 331 TCCTCTGCTATGGGACCTCTACTGTTTGGAACTCGCTACAGACCTCGAAGTTACC 390  
 QY 102 AspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaValGly 121  
 DB 391 CAGGCAATGAAAGATGCCCTGACTCAGCAATATATATGCTATGCCCCATCCATCGGC 450  
 QY 122 LeuProAlaAlaArgSerAlaValAlaGlu-----HisLeuSerGlnGlyValProTyr 139  
 DB 451 TTGCTATCCAGTCGGAGAGAGATTGCTTCTTATACCACTGCTCCCTGGAGGCA----- 501  
 QY 140 LysLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 159  
 DB 502 CCCCTAGAGCTAAGACGCTATTCTGACAGTGCCTCAGCAAGCTATTGACCTTTGT 561  
 QY 160 IleProValLeuAlaGlnThrAlaGlyAlaAlaIleLeuLeuProArgProGlyTyrPro 179  
 DB 562 ITAGCTGTGTGGCC---AACCCAGGCAAAACATCTGGTTCCTCAAGACCTGGTTCTCT 618  
 QY 180 AsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisPheAspLeuIlePro 199  
 DB 619 CTCACAGACTCTGGCTGAGTCTATGGAAATGAGTCAAACTCTCAATTTGTGCA 678  
 QY 200 AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 219  
 DB 679 GAGAAATCTTGGAAATGACCTGAAACAACTGGAATATCTAATTGATGAAAGACAGCT 738  
 QY 220 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 239  
 DB 739 TGCTTCATGTCAATAATCCATCAACCCCTGTGGTTCAGTGTTTCAGCAAAACGTCATCTT 798  
 QY 240 AlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspGluValTyr 259  
 DB 799 CAGAAGATTCTGGCAGTGGCTGCAGGCGAGTGTCTCCCATCTTAGCTGATGAGATCTAT 859  
 QY 260 GlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAla 279  
 DB 859 GGAGACATGGTGTTCGGATTGCAATATGAACCACTGGCCACCTCAGCACCGATGTC 918  
 QY 280 ProValLeuSerIleGlySerLeuSerLysSerThrIleValProGlyTyrArgLeuGly 299  
 DB 919 CCCATCTGCTCTGTGGAGGCTGCCAAGCGCTGGCTGGTTCCTGTGCTGAGGTGGGC 978  
 QY 300 TrpValAlaValTyrAspProThrLysIleLeuLys-----Thr 313  
 DB 979 TGGATCTCTCATCATGACCGAAGACATTTTGGCAATGAGATCCGAGATGGCGCTGGTG 1038  
 QY 314 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheVal 333  
 DB 1039 AAGCTGAGTCAGCGCATTTTGGGA-----CCCTGTACCATTTGTC 1077  
 QY 334 GlnGluAlaLeuProLysIleLeuGluAsnThrLysAlaAspPhePheLysArgIleIle 353

DB 1078 CAGGAGCT----- 1086  
 QY 354 GlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGluAsnLysTyrIle 373  
 DB 1087 ---CTGCTCAAGTCCAATGCTGATCTCTGTATGGGCGTGTGGCTGCCATCCCTGGACTC 1143  
 QY 374 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 393  
 DB 1144 CGG---CCAGTCCGCCCTTCTGGGGCTATGTACCTCATGTTGGAATGAGATGAACAT 1200  
 QY 394 LeuGluGluIleHisAspAspIleAspPheCysCysLysLeuAlaLysGluGluSerVal 413  
 DB 1201 TTCCAGAAATTGAGAACGATGTGGAGTTCACGAGCGGTAGTTGTGAGCAGCTCTGTC 1260  
 QY 414 IleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPheAlaCys 433  
 DB 1261 CACTGCCTCCAGCAACGCTGTGAGTACCGAATTCATCCGAGTGGTCATCACAGTC 1320  
 QY 434 ValProSerSerLeuGlnAspGlyLeuGluArgValLysSerPheCysGlnArgAsn 452  
 DB 1321 CCCGAGTGATGATGCTGGAGCGGTGCAGCGGATCCAGGAGTCTCTGTGAGCAGCAC 1377

# RESULT 8

US-10-198-846-8896/c  
 ; Sequence 8896, Application US/10198846  
 ; Publication No. US20030099974A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lilie, James  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Wang, Youzhen  
 ; APPLICANT: Steimann, Kathleen  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
 ; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
 ; FILE REFERENCE: MRI-049  
 ; CURRENT APPLICATION NUMBER: US/10/198,846  
 ; CURRENT FILING DATE: 2002-07-18  
 ; PRIOR APPLICATION NUMBER: 60/306,220  
 ; PRIOR FILING DATE: 2001-07-18  
 ; NUMBER OF SEQ ID NOS: 14084  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8896  
 ; LENGTH: 870  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 24, 687, 693, 705, 723, 757, 768, 814, 826  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-10-198-846-8896

Alignment Scores:  
 Pred. No.: 4,1e-35 Length: 870  
 Score: 390.00 Matches: 100  
 Percent Similarity: 51.75% Conservative: 48  
 Best Local Similarity: 34.97% Mismatches: 102  
 Query Match: 16.32% Indels: 37  
 DB: 14 Gaps: 7

US-10-019-783-1 (1-461) x US-10-198-846-8896 (1-870)

QY 74 GlusSerGlyProArgPro-----ValLeuProLeuAlaHisGly----- 86  
 DB 849 GAACCAATCCCAACCAAAANAAATGTTTCCNTGGCCCAATGGGGGACCCCTACC 790  
 QY 87 -----AspProSerValPheProAlaPheArgThrAla 97  
 DB 789 TGGTGTGGGAAAACCTGCGCNTACAGACCCNTTAAAGTTTCCC----- 745  
 QY 98 ValGluAlaGluAspAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 117  
 DB 744 ---CAGGCCAAATAAAGATGCCNTGGNACTCGGGGCAAAANAAAGGGCTANGGCC 688

```
QY 118 AlaGlyValGlyLeuProAlaAlaAraSerAlaValAlaGluHisLeuSerGlnGlyVal 137
DB 697 NATCAATCGGGCTTCCCTA-TCCAAATCGGGAAGAAGATTGTTCTTTATTACCCTGT 629
QY 138 ProTyrLys-----LeuSerAlaAspValPheLeuThr-AlaGlyGlyThrGlnAl 155
DB 628 CTTGAAGGGACCCCTAGAGCTAAGGAGCTCATTTGACAAAGTGGCTGCAGCAAGC 569
QY 155 aileuValle-IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProA 175
DB 568 TATTGACCTTTTGTAGCTGTGTGCC---AACCCAGGGCCAAACATCTCGTCCAA 512
QY 175 rgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnLysLeuGluValArgHisP 195
DB 511 GACCTGGTTCCTCTCTACAGACTCTGGCTGAGTCTATGGGAATTGAGGTCAACTCT 452
QY 195 heAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAla 215
DB 451 ACAATTCTGCGAGAGAAATCTTGGGAATTGACCTGAAACAACTGGAATATCTAATTG 392
QY 215 spLysAsnThrThraAlaMetValIleLeuAsnProAsnAsnProCysGlySerValTyrS 235
DB 391 ATGAAGAAGACAGCTGTCTCATTTGCAATAATCCATCAACCCCTGTGGGTGAGTGITCA 332
QY 235 exTyrAspHisLeuAlaLysValAlaGluValAlaArgLysLeuGlyIleLeuValIleA 255
DB 331 GCAACAGTCATCTTCAGAGATTCTGACAGTGGCTGCAGCGCATGTGTGCCCATCTTAG 272
QY 255 laAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValP 275
DB 271 CTGATGAGATCATGAGACATGTTGTTTCCGATTGCAATATGAAACCATCTGCGCCGCC 212
QY 275 heGlyHisIleAlaProValLeuSerIleGlySerLeuSerLysSerTrpIleValProG 295
DB 211 TCAGCAGCATGTCCTCCATCTCTGCTGCTGAGGGCTCGCCAGGGCTGGTGTCTGT 152
QY 295 lyTrpArgLeuGlyTrpValAlaValAlaValTyrAspProThrLysIleLeuLutys----- 312
DB 151 GCTGAGGTTGGCTGGATCTCTCATTCATGACCGAAGAGACATTTTGGCAATGAGATCC 92
QY 313 -----ThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspP 329
DB 91 GAGATGGGCTGTGAGCTGAGCTGAGCGCATT-----TTGGACCCCTGTACCTCGG 41
QY 329 roAlaThrPheVal 333
DB 40 CCGCCACCGCGGTG 27

RESULT 9
US-09-905-173-39
; Sequence 39, Application US/09905173
; Patent No. US20020132295A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay M.
; APPLICANT: WARREN, Patrick V.
; APPLICANT: SWANSON, Ronald V.
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: ENZYMES HAVING TRANSAMINASE AND AMINOTRANSFERASE ACTIVITY AND MET
; FILE OF INVENTION: US THEREOF
; FILE REFERENCE: DIVER1240-7
; CURRENT APPLICATION NUMBER: US/09/905,173
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 09/412,184
; PRIOR FILING DATE: 1999-10-04
; PRIOR APPLICATION NUMBER: US 09/389,537
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: US 08/646,590
; PRIOR FILING DATE: 1996-05-08
; PRIOR APPLICATION NUMBER: US 08/599,171
; PRIOR FILING DATE: 1996-02-09
; PRIOR APPLICATION NUMBER: US 09/481,733

; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: US 09/069,226
; PRIOR FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 1185
; TYPE: DNA
; ORGANISM: Aquifex
US-09-905-173-39

Alignment Scores:
Pred. No.: 1,35e-31 Length: 1185
Score: 361.50 Matches: 109
Percent Similarity: 46.57% Conservative: 81
Best Local Similarity: 26.72% Mismatches: 173
Query Match: 15.13% Indels: 45
DB: 10 Gaps: 14

US-10-019-783-1 (1-461) x US-09-905-173-39 (1-1185)
QY 64 IleArgTyrLysIleSerAlaSerValGluGluSerGlyProArgProValLeuProLeu 83
DB 1 ATCAGAAAGAGGACTTGCAGTAGGCTAAGTACCTTAAACCTTCCCCCAACGCTGACCATA 60
QY 84 AlaHis-----GlyAspProSerVal 90
DB 61 ACCGCAAAAGCAAAAGAAATTAAGGCTTAAGGAGTGGAGCTTATAGGTTTGGAGCGGA 120
QY 91 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaLeuArgThr 110
DB 121 GRACCTGACTTCGACACACCCGACTTCATAAGGAAGGCTGTATAGGCTTTAAGGGA 180
QY 111 GlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 130
DB 181 GGAAGACCAAG---TACGCTCCCTCCGCGGAATACCAGAGCTCAGAGAGCTATAGCT 237
QY 131 GluHisLeu-----SerGlnGlyValProTyrLysLeuSerAlaAspValPheLeu 148
DB 238 GAAAACTACTGAAAGAAACAAAGTTGAGTACAAACCTTCA-----GAGATAGCTGT 291
QY 149 ThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGly 168
DB 292 TCCGCGAGGAGCGAAATGTTCTCTCTCATATTATCATGCTATATGAGC---GAAGGA 348
QY 169 AlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsn 188
DB 349 GACGAGGTTTACTACCTAGCCCTTACTGGTAACCTTACCCCGAA-----AspLys 201
QY 189 LysLeuGluValArgHisPheAspLeuIlePro-----AspLys 201
DB 394 -----CAGATAAGTTCTTCGAGGGGTTCCTGTTAGGTTCTCTAAAGAAAGAGAAA 447
QY 202 GlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaVal 221
DB 448 GGATTTCAATTAAGTCTGGAAGATGGAAGAAAGAGTTACGAGAGAGAGAAAGCTATA 507
QY 222 ValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuAlaLys 241
DB 508 GTCAATACTCTCCGACCAACCCCTGCTGCTGTTTACGAAGAGGAGGAACTTAAGAAA 567
QY 242 ValAlaGluValAlaArgLysLeuGlyIleLeuValIleAlaAspValTyrGlyLys 261
DB 568 ATAGCGGAGTTTTCGCTGGAGAGGGGCATTTTCAATAATTTCCGATGAGTGTAGTAC 627
QY 262 LeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleAlaPro--- 280
DB 628 TTGCTTTACGCTGATGCAAAATTTGTAGCCCTGCTCTTCTCGGATGAGTAAAGAAC 687
QY 281 ---ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGly 299
DB 688 ATAACCTTCACGGTAAACGCTTTTTCGAAAGAGCTATTTCATGACTGCTGTTGGCAATAGGT 747
```





Db	1036	ATCCGCGCAAGCGCCCGCAGGACACCGTCCAGCTGCGCGCGTGTATCTCGGAGGAGCC	1095
Qy	412	SerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArgIleThrPhe	431
Db	1096	GAGGTCCGGTCTGTCCTCCCGGAGGCGCTTCGGCACCGCGGCGTATCTGCGGCTGTCTGTAC	1155
Qy	432	AlaCysValProSerSerLeuGlnAspGlyLeuGluArgVal-LysSerPheCysGlnAr	451
Db	1156	GCCCTGGGTGACGAGGATCTGTTCGAGGCGCTGACCGCGATTCAAGAGCTGCTGTCCGAG	1215
Qy	451	gAsnLys	453
Db	1216	GCGAAGG	1222
RESULT 11			
US-10-156-761-1/c			
; Sequence 1, Application US/10156761			
; Publication No. US20030119018A1			
; GENERAL INFORMATION:			
; APPLICANT: OMURA, SATOSHI			
; APPLICANT: IKEDA, HARUO			
; APPLICANT: ISHIKAWA, JUN			
; APPLICANT: HORIKAWA, HIROSHI			
; APPLICANT: SHIBA, TADAYOSHI			
; APPLICANT: SAKAKI, YOSHIYUKI			
; APPLICANT: HATTORI, MASAHIRA			
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES			
; FILE REFERENCE: 249-262			
; CURRENT APPLICATION NUMBER: US/10/156,761			
; CURRENT FILING DATE: 2002-05-29			
; PRIOR APPLICATION NUMBER: JP 2001-204089			
; PRIOR FILING DATE: 2001-05-30			
; PRIOR APPLICATION NUMBER: JP 2001-272697			
; PRIOR FILING DATE: 2001-08-02			
; NUMBER OF SEQ ID NOS: 15109			
; SEQ ID NO 1			
; LENGTH: 9025608			
; TYPE: DNA			
; ORGANISM: Streptomyces avermitilis			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (4187715)			
; OTHER INFORMATION: a, t, c, g, other or unknown			
US-10-156-761-1			
Alignment Scores:			
Pred. No.:	1,08e-24	Length:	9025608
Score:	355.50	Matches:	104
Percent Similarity:	47.04%	Conservative:	95
Best Local Similarity:	24.59%	Mismatches:	178
Query Match:	14.87%	Indels:	46
DB:	14	Gaps:	15
US-10-019-783-1 (1-461) x US-10-156-761-1 (1-9025608)			
Qy	51	IleLeuAlaThrThrGlyAlaLysAsnSerIleArgAlaIleArgTyrLysIleSerAla	70
Db	5964250	GTCCTCCGCGCGAGTCGCGCGGATCTCCGAGTCGCGCCACCTCCGCGTGTATCTCGGAGGAGCC	5964191
Qy	71	SerValGluGluSerGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal	90
Db	5964190	AAGGCCCTCAAGCCCGCGGAGCTGCGGTGATCGGCTTCGCGCGCGGTGAGCCGAC---	5964134
Qy	91	PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaLeuArgThr	110
Db	5964133	-----TTCCGACCCCGGACTACATCGTCCAGCGCGCATCGAGGCTGTCTCGAAC	5964083
Qy	111	GlyGlnPheAsnCysTyrAlaAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla	130
Db	5964082	CCGAAGTACCACCGCTACACCGCGCGCGGCTGCTCCCGAGTGAAGCGCCCATCGCC	5964023
Qy	131	GluHisLeuSerGlnGlyValProTyrLysLeuSerAlaAspValPheLeuThrAla	150



[illegible]

DB: 10 Gaps: 0

US-10-019-783-1 (1-461) x US-09-878-574-2935 (1-370)

QY 274 ValPheGlyHisIleAlaProValIleuSerIleGlySerLeuSerIlySerTrpIleVal 293  
 Db 1 GTTTTGGCTATATGTTCTGACTCCTTGGCTCATTTTCTAAGAGATGGATGTT 60

QY 294 ProGlyTrpArgLeuGlyTrpValAlaValTyPProThrIlySileLeuGlyIlyThr 313  
 Db 61 CCTGGATGGAGGCTTGGTTGTGTGACAAATGATCCATCTCGGCACCTTTTGAAGAAATCCA 120

QY 314 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheVal 333  
 Db 121 AAGTAGATAGAGCGCTTAAAGAGTACTTGTATGCTGGTGGAGGTCCAGCACCTTCATC 180

QY 334 GlnGluAlaLeuProIlySileLeuGluAsnThrLysAlaAspPhePheIleIle 353  
 Db 181 CAGGACGCTGTACCTCAGATAATTGAGCATCTGAAAAGGTTTCTTCAAGAAATCAAT 240

QY 354 GlyLeuLeuIlySileGluSerSerGluIleCysTyrArgGluIleIlySileGluAsnIlyThr 373  
 Db 241 GATAAATTGAGGATGTTGACATATATGTGTAAAGAGATTGAAGATATTCATACAT 300

QY 374 ThrCysProHisIlySileProGluGlySerMetPheValMetValIlySileGluAsnHisLeu 393  
 Db 301 ATATGCCCTTATAAACAGAAAGGTCATGGCTATGATGTGAACTTAACCTTTCATCT 360

QY 394 LeuGluGlu 396  
 Db 361 CTGAGGAT 369

RESULT 15

US-10-156-761-1

; Sequence 1, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 1

; LENGTH: 9025608

; TYPE: DNA

; ORGANISM: Streptomyces avermitilis

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (4187715)

; OTHER INFORMATION: a, t, c, g, other or unknown

US-10-156-761-1

Alignment Scores:

Pred. No.: 5,51e-21 Length: 9025608

Score: 323.50 Matches: 123

Percent Similarity: 44.86% Conservative: 82

Best Local Similarity: 26.80% Mismatches: 197

Query Match: 13.54% Indels: 58

DB: 14 Gaps: 18

US-10-019-783-1 (1-461) x US-10-156-761-1 (1-9025608)

US-10-019-783-1 (1-461) x US-10-156-761-1 (1-9025608)

Mon Nov 10 11:53:06 2003

```
Db 2449104 ATCCAGGCGCGCTCGGCGCGCCAGTCCATCCAGGAGCTGACCGCACCGGGC----- 2449157
Qy 350 LysArgIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrArgGluIleLysGlu 369
Db 2449158 -----GGCGGCTGTACGAGCAGCGCGACCGCGCTGGCAGAGCTCAACGAG 2449205
Qy 370 AsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeu 389
Db 2449206 ATCCGGGCGTCTCGTGGTG--AAGCCGAAGGGCGGTTGTACGGCTTCCCGCGCATC 2449262
Qy 390 AsnLeuHisLeuLeuGluGluIleHisAspAspIleAspPheCysLysLeuAlaLys 409
Db 2449263 GAC---CCCACCGTGSCACAGATCCACGACGAGAGATTCTGCTCTGGACCTCTGCTG 2449319
Qy 410 GluGluSerValIleLeuCysProGlySerValLeuGlyMetGluAsnTrpValArg 428
Db 2449320 CGGGAGAGATCCAGGTGTCAGGGCACCGGCTTC-----AACTGGCCCCGC 2449367
```

Search completed: November 9, 2003, 09:07:34  
Job time : 17729.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: November 8, 2003, 19:50:55 ; Search time 3357.72 Seconds  
(without alignments)  
3988.346 Million cell updates/sec

Title: US-10-019-783-2  
Perfect score: 2856  
Sequence: 1 MARVQSDGVAAAGLAFAA.....LGRKSFQCKRKNSSDDC 551

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Maximum Match 0%  
Listing first 45 summaries

Command line parameters: -MODEL=frame\_plus.p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10019783/runat\_07112003\_161643\_19848/app\_query.fasta\_1.1358  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10019783 @CGN 1.1 4125 @runat\_07112003\_161643\_19848 -NCPU=6 -ICPU=3  
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estio.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mus.\*  
23: em\_gss\_man.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1331.5	46.6	1143	11	AY104359	AY104359 Zea mays
2	1214	42.5	770	10	BF268125	BF268125 HV_Cra001
3	1178.5	41.3	813	14	CB625993	CB625993 OSIIla015r
4	1114	39.0	657	14	CA015631	CA015631 HT14015r
5	1026.5	35.9	837	14	CB668395	CB668395 OSJNEd16D
6	1012.5	35.5	891	14	CD437584	CD437584 ELO1N0502
7	1006	35.2	606	13	BQ466157	BQ466157 HT01L21T
8	1000	35.0	843	14	CB668396	CB668396 OSJNEd16D
9	1000	35.0	878	14	CB643965	CB643965 OSJNEd16D
10	1000	35.0	902	14	CB618032	CB618032 OSIIla01U
11	999	35.0	849	14	CB671005	CB671005 OSJNEd04F
12	986.5	34.9	809	14	CB618032	CB618032 OSIIla01U
13	995	34.8	832	14	CB652840	CB652840 OSJNEC03E
14	981	34.3	883	14	CB646007	CB646007 OSJNEB08C
15	966.5	33.8	792	14	CB670913	CB670913 OSJNEd04D
16	966	33.8	565	10	BF065437	BF065437 HV_CBS001
17	954	33.4	619	10	EG313019	EG313019 WHE0091_D
18	921	32.2	585	14	CA599545	CA599545 waw1C.pko
19	918	32.1	836	14	CB670014	CB670014 OSJNEd02C
20	899.5	31.5	745	14	CB678552	CB678552 OSJNEd16L
21	891.5	31.2	751	14	CB668540	CB668540 OSJNEd16H
22	883.5	30.9	740	14	CB646006	CB646006 OSJNEd16C
23	874	30.6	816	14	CB668541	CB668541 OSJNEd16H
24	854	29.9	577	13	BQ663099	BQ663099 HT01L21w
25	854	29.9	743	14	CB675132	CB675132 OSJNEd10K
26	845.5	29.6	604	9	AW257959	AW257959 687064G08
27	839.5	29.4	811	14	CB671004	CB671004 OSJNEd04F
28	832.5	29.1	806	14	CB643964	CB643964 OSJNEB04P
29	828.5	29.0	577	10	BE230084	BE230084 99A3319_R
30	811.5	28.4	806	14	CD439867	CD439867 ELO1N0530
31	809	28.3	668	14	CD230874	CD230874 SSI_20_D1
32	808	28.3	744	14	CB670914	CB670914 OSJNEd04D
33	807	28.3	683	14	CD210710	CD210710 HSI_49_A1
34	806	28.2	763	13	BQ967714	BQ967714 QHB30L20.
35	803	28.1	791	13	BQ969834	BQ969834 QHB39K17.
36	796	27.9	734	14	CA401705	CA401705 ELO1N0424
37	795.5	27.9	596	10	BE365047	BE365047 P11_23_B1
38	794.5	27.8	593	10	BE365050	BE365050 P11_23_A1
39	793	27.8	777	13	BQ915490	BQ915490 QHB14M12.
40	790.5	27.7	562	13	BU050754	BU050754 1111033B0
41	790.5	27.7	589	10	BE365012	BE365012 P11_23_A1
42	783	27.4	473	13	BQ469047	BQ469047 HM03E06r
43	783	27.4	734	13	BQ970717	BQ970717 QHB42014.
44	782	27.4	481	10	BE200495	BE200495 WHE2258_G
45	781	27.3	502	10	BE587861	BE587861 WHE0670_D

ALIGNMENTS

RESULT 1	AY104359	Zea mays	PC0115235 mRNA sequence.	1143 bp	mRNA	linear	HTC 16-OCT-2002
LOCUS	AY104359						
DEFINITION	Zea mays						
ACCESSION	AY104359						
VERSION	AY104359.1		GI:21207437				
KEYWORDS	HTC.						
SOURCE	Zea mays						
ORGANISM	Zea mays						
REFERENCE	1	(bases 1 to 1143)					

AUTHORS		Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.	
TITLE		Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes	
JOURNAL		Unpublished (2002)	
REFERENCE		2 (bases 1 to 1143)	
AUTHORS		Coe, E.H.	
JOURNAL		Direct Submission	
TITLE		Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA	
COMMENT		If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, maizeimap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.	
FEATURES		Location/Qualifiers	
source		1..1143	
		/organism="Zea mays"	
		/mol_type="mRNA"	
		/db_xref="MaizeDB:636702"	
		/db_xref="taxon:4577"	
		/clone_lib="Maize Mapping Project/DuPont Consensus Library"	
		/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BAGs in conjunction with the Maize Mapping Project"	
BASE COUNT		314 a	267 c 280 g 279 t 3 others
ORIGIN			
Alignment Scores:			
Pred. No.:		2,066-123	Length: 1143
Score:		1331.50	Matches: 256
Percent Similarity:		86.93%	Conservative: 30
Best Local Similarity:		77.81%	Mismatches: 42
Query Match:		46.62%	Indels: 2
DB:		11	Gaps: 1
US-10-019-783-2 (1-551) x AY104359 (1-1143)			
QY	218	AlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPhe	237
DB	14	GGCGTGGACGACACTGTCGAGCGACCTTCATACAAAGCTGACGACCGACATCTTC	73
QY	238	LeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAla	257
DB	74	GTCAACCGCGGAGN-ACGCAAGCCATCGAGGTGGTGTCTCAGTCTCGCGCAG-CCG	129
QY	258	GlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe	277
DB	130	GGACCAACATATCTGCTCCGAGCGGGCTATCCGACTACGAGCGCGCGAGGGCTG	189
QY	278	AsnArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIle	297
DB	190	CACAACTCGAAGTTCGCGGTTCAATCTGATCCCGAGAGAGGGTGGAGATTGACATC	249
QY	298	AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsn	317
DB	250	GACGGTCTGGAGTCGATCCCGACAGAACACACCGCCATGCTATATAACCCCAAC	309
QY	318	AsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLys	337
DB	310	AACCTTTCGGGAGTCTCTACACCGAGAGCATTTGGCCAAAGTCCGCGAGGTGCAAGG	369
QY	338	ArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAla	357
DB	370	AAGCTTGGAAATCTGTCATCGCTGATGAGGTGTACGGAATCTGTTGGGACACC	429
QY	358	ProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeu	377
Db		CCTTTCGTCCTCCACGGGTGTCTTTGGCCACATTGCCCTGTGTGACCATAGGATCACTG	489
QY		SerLysSerTrpIleValProGlyTyrArgLeuGlyTrpValAlaValTyrAspProArg	397
Db		TCCAAAGATGATAGTAGTGCCTGGGTGGCGACTTGGTTGGGTGGCGGTATGTATGCCAAC	549
QY		LysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr	417
Db		AAGTTCCTGCAGAAACCAAGATCATTCATCGATCACAAACTTCCCTTAACGTTTCAACA	609
QY		AspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGlnAsnThrLysGluAsp	437
Db		GATCCGGCAACTTTTGTTCAGGAGCGCTTCCCATATTTCTTAAGAACACAAAGAAAGAT	669
QY		PhePheLysAlaIleIleGlyLeuLysGluSerSerGluIleCysTyrLysGlnIle	457
Db		TTTTTCAAGAGATCATTTGGTCTGCTAGCGAAACATCAGAGATATGTTTCAGCGGAATA	729
QY		LysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetVal	477
Db		AAGGACATCAAGTGCATCACATGCCCTCACAAAGCCAGAGGTTCCATGTTTGTTCATGCTG	789
QY		LysLeuAsnLeuHisLeuLeuGluIleAspAspIleAspPheCysCysLysLeu	497
Db		AAACTAAATTTGTATCTTTTGGAGAGCATCCATGATGATATGATTTTGTTCGAGCTG	849
QY		AlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpVal	517
Db		GCAAAAGAGAGTCCGTGATTTTGTTCAGGAGTGTTCGGAATGGAATACTGGATC	909
QY		ArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSer	537
Db		CGTATCATCTTCGCCATTGATTCATCTCTCTCTTGTGTCCTTTGAGAGGCTGAAATCT	969
QY		PheCysGlnArgAsnLysLysArgAsn	546
Db		TTCTGCCAAAGGCATAAGAGAGAGAT	996
RESULT 2			
LOCUS		BF268125 770 bp mRNA linear EST 23-OCT-2001	
DEFINITION		HV_CEA0019P10f Hordeum vulgare seedling green leaf EST library	
ACCESSION		HVCDNA0004 (Blumeria challenged) Hordeum vulgare subsp. vulgare	
VERSION		CDNA clone HV_CEA0019P10f, mRNA sequence.	
KEYWORDS		BF268125 GI:13263792	
SOURCE		EST.	
ORGANISM		Hordeum vulgare subsp. vulgare	
REFERENCE		1 (bases 1 to 770)	
AUTHORS		Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Wei, F., Begum, D., Fisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton, R.D., Oates, R., and Main, D.	
TITLE		Development of a genetically and physically anchored EST resource for barley genomics: Blumeria infected incompatible (Mla13) seedling leaf cDNA library	
JOURNAL		Unpublished	
COMMENT		On Nov 17, 2000 this sequence version replaced gi:11199120.	
CONTACT		Contact: Wing RA	
INSTITUTION		Clemson University Genomics Institute	
ADDRESS		100 Jordan Hall, Clemson, SC 29634, USA	
TELEPHONE		Tel: 864 656 7288	
FAX		Fax: 864 656 4293	
EMAIL		Email: rwing@clemson.edu	
TOTAL HQ BASES		Total hg bases = 586	
SEQ PRIMER		Seq primer: AATTAACCTCTCACTAAAGG	
HIGH QUALITY		High quality sequence stop: 611.	
LOCATION/QUALIFIERS			



```

source      1. 770
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="C116155 (M1a13)"
/db_xref="taxon:112509"
/clone="HV_CEA0019P10f"
/tissue_type="seedling green leaf"
/lab_host="RUC121"
/library="Hordeum vulgare seedling green leaf EST"
/library_lib="HVCNRA0004 (Blumeria challenged)"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI; C.I. 16155 (M1a13) plants were greenhouse grown in the R Wise lab at Iowa State University, Ames, IA; 7 day old green seedlings were challenged with isolate A27 (AvrMla13) of Blumeria graminis f. sp. hordei, and leaves were harvested 20 and 24 hr post-inoculation and snap frozen; uninoculated leaves were harvested 20 hr post-inoculation (Wei, Wise). In the TJ Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one cDNA library was made, and 1 million pfu were in vivo excised to give pBlueScript SK(-) cDNA phagemids (Choi, Close). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinohs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/gpages/bgn/31/cover.html)"
BASE COUNT 164 a 247 c 216 g 143 t
ORIGIN
Alignment Scores:
Pred. No.: 7,83e-112 Length: 770
Score: 1214.00 Matches: 241
Percent Similarity: 96.11% Conservative: 6
Best Local Similarity: 93.77% Mismatches: 8
Query Match: 42.51% Indels: 2
DB: 10 Gaps: 0

US-10-019-783-2 (1-551) x BF268125 (1-770)
QY 183 AlapheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGln 202
Db 3 GCCTTCGACGACGCGCTCGNGCGCGAGAGCGCGTCCGCCCGCGCTGGGACCGGCGAG 62
QY 203 PheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaAlaArgSerAlaValAlaGluHis 222
Db 63 TTCAACTGTCTACCCCGCGCGCTCGCGCTCCCGCGCGACGACGAGCGCGCTGGCAGAGCAC 122
QY 223 LeuSerGlnGlyValProTyrMetLeuSerAlaAspAspValPheLeuThrAlaGlyGly 242
Db 123 CTGTGCGAGGCGGTGCGGTACAGGCTATCGCGGACGACGTCTTCTCACCAGCGCGCGGG 182
QY 243 ThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeu 262
Db 183 ACCCAGCGCATCGAGGTCTAATCCCGGTGCTGGCCACGACGCGCGCGGCCAACATTCGTG 242
QY 263 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluVal 282
Db 243 CTCCCCAGGCGAGGCTATCCAACTACGAGGCGCGCGCGCTTCAACAGGCTGGAGGTC 302
QY 283 ArgHisPheAspLeuIleProAspLysGlyTyrPheGluAlaAspSerLeuGluSer 302

```

```

Db 303 CGGCATTTGACCTCATCCCGGACAGAGGGTGGGAGATCGACATCGACTCGCTGGAATCC 362
QY 303 IleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySer 322
Db 363 ATCGCGGACAGAACACACCGCATGTGTATATAAACCCCAACACCCGCTGGCGGAGC 422
QY 323 ValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeu 342
Db 423 GTTTACTCTCAGACCATCTGTCCAAAGTTCGGAGGTGGGAGGCTCGGAATATTG 482
QY 343 ValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMet 362
Db 483 GTGATTGCTGACAGGTATACGGCAAGCTGTCTTGGCGACGCGCGCTTCATCCCAATG 542
QY 363 GlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTrpIle 382
Db 543 GGAGTGTTGGGCATCATCCCTGTGTCTCCATAGGCTCTCTGTCTTAACTATGTGATA 602
QY 383 ValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLys-IleLeuGlnGly 402
Db 603 GTGCCTGGATGGCGCTTGGATGGGTAGCGGTGTACGAAACCCCAAGGATCTTCGCGGA 662
QY 402 UnhryIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPh 422
Db 663 AACTAAGAACTTTTACATTCATTCAGTAATCTTAAATCTTCGACAGCCCAACACTT 722
QY 422 IleGlnAlaAlaLeuProGln-IleLeuGluAsnThrLysGluAsp 437
Db 723 TATTCAGGCGCATCTTCTCAGATTTTGGAAACACCAAGGAGGAGAT 769

RESULT 3
CB625993
LOCUS      813 bp mRNA linear EST 08-APR-2003
DEFINITION OSIIIEa15015.f OSIIIEa Oryza sativa (indica cultivar-group) cDNA
clone OSIIIEa15015 5', mRNA sequence.
ACCESSION CB625993
VERSION CB625993.1 GI:29620982
KEYWORDS EST
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 813)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088 USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: 0 column: 15
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1. .813
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIIEa15015"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIIEa"

```

/note=vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
XhoI; Lesion Mimic SPL 11"

BASE COUNT 194 a 235 c 191 g 193 t  
ORIGIN

## Alignment Scores:

Pred. No.: 3,21e-108 Length: 813  
Score: 1178.50 Matches: 220  
Percent Similarity: 91.14% Conservative: 23  
Best Local Similarity: 81.18% Mismatches: 1  
Query Match: 41.26% Indels: 1  
DB: 14 Gaps: 1

US-10-019-783-2 (1-551) x CB625993 (1-813)

QY 158 ILeSerAlaSerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAsp 177  
Db 1 ATCAGCGCCAGGCGCCACTACCGCGCGCGCGCGCGCGCTCTCGCGCTCGCCACGGGAC 60  
QY 178 ProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaAlaAla 197  
Db 61 CCTTCGGTTCGCCAGCTTCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120  
QY 198 ValArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSer 217  
Db 121 CTCGGCTCCGCGACTTCATCTGCTACCCCGCGCGCGCGCGCGCGCGCGCGCGCG 180  
QY 218 AlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPhe 237  
Db 181 GCTGTGGCAGATCATTTGTCACGCGACCTCCCATACAGCTATCTTCTGATGACATCTTC 240  
QY 238 LeuThrAlaGlyGlyThrGluAlaLeuGluValLeuLeuProValLeuAlaGlnThrAla 257  
Db 241 CTAACCGCTGGAGGAACCTCAGCGCATCGAGTGTGTAATCTCAATCTTGCCCAA---CCT 297  
QY 258 GlyAlaAsnLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPhe 277  
Db 298 GGCACAAACATATTGCTTCTTAGACGAGCTACCCAACTATGAAGCTCGAGCGCGCTTC 357  
QY 278 AsnArgLeuGluValArgHisPheAspLeuLeuProAspLysGlyTyrGluLeuAspIle 297  
Db 358 AACAACTTGAAGTTCGTCATTTGATCTTATCTGAGAAAGCGCTGGAGATTGACCTT 417  
QY 298 AspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleLeuAsnProAsn 317  
Db 418 AACTCCCTAGATCATTTGCGGACAAAGAACACTACTGCGTAGTCAATCAATCCCAAT 477  
QY 318 AsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLys 337  
Db 478 AATCCATCGCGGAATGTGTACATTACGAGCATTTATCCAAAGTGGCGAGGTAGCAAGG 537  
QY 338 ArgLeuGlyIleLeuValIleAlaAspLeuValTyrGlyLysLeuValLeuGlySerAla 357  
Db 538 AAGCTTGGANATTGGTAATTACTGATGAGGTGTAAGTAATTGTTTGGAGATTCC 597  
QY 358 ProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeu 377  
Db 598 CCAATTGTCCTCAATGGGTGCTTTGGGCACATCGTACCAATATTAAACCATAGGATCGCTA 657  
QY 378 SerLysSerThrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArg 397  
Db 658 TCAAGAGGTGATAGTGCCTGGATGGGACATTGTTGGGTAGCAATATGTGACCCCAAG 717  
QY 398 LysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThr 417  
Db 718 AGACTCTACAGAACCAAGATTGCAACATTAAATTACTAATTCTTCTTAATGTTTCACT 777  
QY 418 AspProAlaThrPheIleGlnAlaLeuPro 428  
Db 778 GATCCAGCAACTTTCATTCCAGGAGCTCTACCG 810

RESULT 4  
CA015631

LOCUS CA015631 657 bp mRNA linear EST 23-OCT-2002  
DEFINITION HT14015r HT Hordeum vulgare subsp. vulgare cDNA clone HT14015  
5-PRIME, mRNA sequence.  
ACCESSION CA015631  
VERSION CA015631.1 GI:24292975  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
; Triticeae; Hordeum.  
REFERENCE 1 (bases 1 to 657)  
AUTHORS Zhang,H., Potokina,E., Michalek,W., Weschke,W., Stein,N. and Graner  
A.  
Barley ESTs from germinating seeds  
Unpublished  
Contact: Stein Nils  
Molecular Markers Group, Department Genbank  
Institute of Plant Genetics and Crop Plant Research (IPK)  
Corrensstr. 3, 06466, Gatersleben, Germany  
Tel: 039482-5522  
Fax: 039482-5595  
Email: stein@ipk-gatersleben.de  
Insert Length: 657 Std Error: 0.00  
Plate: 14 row: 0 column: 15  
Seq primer: M13rev.  
Location/Qualifiers  
1. 657  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="barke"  
/db\_xref="GABI:264169"  
/db\_xref="taxon:112509"  
/clone="HT14015"  
/tissue\_type="endosperm early"  
/dev\_stage="0-16 hours after imbibition"  
/lab\_host="XL10-Gold"  
/clone\_lib="HT"  
BASE COUNT 140 a 209 c 187 g 121 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 7,34e-102 Length: 657  
Score: 1114.00 Matches: 215  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 39.01% Indels: 0  
DB: 14 Gaps: 0  
US-10-019-783-2 (1-551) x CA015631 (1-657)  
QY 203 PheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHis 222  
Db 12 TTCACTGTACCCCGCGCGCTCGCGCTCCCGCGCACGAAGCGCGCTGGCAGACAC 71  
QY 223 LeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGly 242  
Db 72 CTCTCGAGCGGTGGTACATGCTATCGCGCGAGCGCTTCTCTACCGCGCGCGG 131  
QY 243 ThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThrAlaGlyAlaSerIleLeu 262  
Db 132 ACCAGCGGATCAGGTCATAATCCCGGTGCTGGCCACAGCGCGCGCGCAACATTCG 191  
QY 263 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaPheAsnArgLeuGluVal 282

```

Db 192 CTCGCCAGGCGAGCTACCAAACTACGAGGCGCGCGCGCTTCAACAGGCTGGAGTC 251
Qy 283 ArgHisPheAspLeuLeuProAspHisGlyTyrGluValAspLeuAspSerLeuGluSer 302
Db 252 CGGCATTTCGACCTCATCCCGCAAGAGGCTGGAGATCGACATCGATCGCTCGTGAATCC 311
Qy 303 IleAlaAspLeuAsnThrAlaMetValIleLeuAsnProAsnAsnProCysGlySer 322
Db 312 ATGCCGACAGAACACCCCGCATGTCATCATAAACCCCAACACCCGCTGGCGAGC 371
Qy 323 ValTyrSerTyrAspHisLeuSerHisValAlaGluValAlaLysLeuGlyLeu 342
Db 372 GTTTACTCTTACGACCATCTGTCAAAGTGGCGAGGTGGCGAAAGGCTCGGAATATG 431
Qy 343 ValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheLeuMet 362
Db 432 GTGATTCTGACAGAGGTATACGCAAGCTGGTTCTGGCGAGCGCCCGTTTATCCCAATG 491
Qy 363 GlyValPheGlyHisLeuThrProValLeuSerIleGlySerLeuSerIleThr 382
Db 492 GGAGTGTGTTGGGCACATCACCCCTGTGTCTCATAGGCTCTCTGTCCAAAGTCATGGATA 551
Qy 383 ValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLeuIleGluGlu 402
Db 552 GTCCCTGGATGGCGGCTTGATGGTGGTGGTGTACGACCCCAAGATCTTTACAGGAA 611
Qy 403 ThrLysLeuSerThrSerLeuThrAsnTyrLeuAsnValSerThr 417
Db 612 ACTAAGATCTCTACATCAATTACGAATTACCTCAATGTCTCGACA 656

RESULT 5
CB668395 837 bp mRNA linear EST 09-APR-2003
OSJNEd16D17.f OSJNEd16D17 5', mRNA sequence.
LOCUS OSJNEd16D17 5', mRNA sequence.
DEFINITION CB668395
ACCESSION CB668395
VERSION 1
KEYWORDS CB668395.1 GI:29672120
SOURCE EST.
ORGANISM Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoideae; Oryza.
REFERENCE 1 (bases 1 to 837)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel.: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 16 row: D column: 17
Seq primer: gta aac cga cgg cca gtc.

FEATURES
source
1..837
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEd16D17"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEd"

```

```

/note="Vector: pBluescript II KS +; Site.1: EcoRI; Site.2:
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT 186 a 256 c 237 g 158 t
ORIGIN
Alignment Scores:
Pred. No.: 6,93e-93 Length: 837
Score: 1026.50 Matches: 204
Percent Similarity: 76.38% Conservative: 32
Best Local Similarity: 66.02% Mismatches: 37
Query Match: 35.94% Indels: 36
DB: 14 Gaps: 5
US-10-019-783-2 (1-551) x CB668395 (1-837)
Qy 67 HisAlaGluAlaThrAlaAsnGlyHisGly-GluAlaThrAlaAsnGlyLysThrAsnGlu 86
Db 13 CACCAACAACAACAGACGACGCGCATGGCACCGCAGCAGCGCG----- 55
Qy 86 yHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluHi 106
Db 56 -----GCGGCGCGCGCGAGCAGCAACGCG----- 79
Qy 106 sAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGluGlu 126
Db 80 -----GCGGCGCGAGAGCAGCGC----- 97
Qy 126 uGluGluAlaValGluTyrAsnPhe---AlaGlyAlaLysAspGlyValLeuAlaAlaLath 145
Db 98 ---AGCAGCAAGGAGTGGAGGCTGACGCGCGCGAGGCGCGCGCGGCGGCGGCGGCG 153
Qy 145 rGlyAlaAsnMetSerIleArgAlaIleArgTyrIleSerAlaSerValGluGluLyl 165
Db 154 GGGGGACAAGATGAGCATCCGGCGCGCTCGGTACAGATCAGCGCGCGCGCGCGCGCG 213
Qy 165 sGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheAr 185
Db 214 CGGCGCGCGCGCGCTCGCGCTCGCCACGCGGACCCCTCCGTTCCCGAGTTCGG 273
Qy 185 gThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGluPheAsnCy 205
Db 274 CACCGCGCGCGAGCGCGAGGCGCGCTCGCGACGCGCTCGCGCGCGCTCAACTG 333
Qy 205 sTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerG 225
Db 334 CTACCCCGCGCGCGCTCCCGCGCGCGCGCGCGCTGTGTGGCAGATCATTTGTCAAG 393
Qy 225 nGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGluAl 245
Db 394 CGACCTCCCATACAAGCTATCTTCTGATGACATCTCTTAACCGCTGGAGGAACTCAGC 453
Qy 245 alieGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAlaAsnIleLeuLeuProAr 265
Db 454 CATCGAGGTCTGAATCTCAATCTTGCCTTGCCTAG 510
Qy 265 gProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPh 285
Db 511 ACCAGGCTACCCAACTATGAAGCTCGAGCGCGCTCAACACCTTGAAGTTCGTCACT 570
Qy 285 eAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAs 305
Db 571 TGATCTTATCTCGAAGGCTGGAGATTGACCTTAACCTCCCTAGATCTATTTCGGA 630
Qy 305 pLysAsnThrAlaMetValIleAsnProAsnAsnProCysGlySerValTyrSe 325
Db 631 CAAGAACACTACTGCGATAGTATATAATCCCAATATCCATCGCGGAATGTGTACAC 690
Qy 325 rTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAl 345
Db 691 TTACGAGCATTTATCCAGGTGGAGGTAGTAGCAGGAGGCTTGGATATTGGTAATTAC 750
Qy 345 aAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheLeuProMetGlyValPh 365

```

Db 751 TGATGAGGTATGGCAATTTGGTTTTCGGAGTTCCCATTTGTCCTCCCATGGGTGCTT 810

QY 365 eGlyHisIleThrProValLeuSer 373  
 Db 811 TGGCACATCGTACCAATTAACC 835

RESULT 6  
 CD437584  
 LOCUS CD437584 891 bp mRNA linear EST 03-JUN-2003  
 DEFINITION EL01N0502F06.b Endosperm\_5 Zea mays cDNA, mRNA sequence.  
 ACCESSION CD437584  
 VERSION CD437584.1 GI:31353227  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 891)  
 Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and  
 Messing, J.  
 TITLE Sequencing of the maize endosperm ESTs  
 JOURNAL Unpublished  
 COMMENT Dr. Joachim Messing's lab  
 Waksman Institute, Rutgers University  
 130 Frelinghuysen Rd., Piscataway, NJ 08854, USA  
 Tel: 732-445-3801  
 Fax: 732-445-5735  
 Email: jlai@waksman.rutgers.edu  
 Seq primer: T3.  
 FEATURES  
 source  
 1..891  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultivar="W22"  
 /db\_xref="taxon:4577"  
 /tissue\_type="Endosperm of 7-23DAP"  
 /clone\_lib="Endosperm 5"  
 /notes="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 170 a 294 c 259 g 168 t

ORIGIN

Alignment Scores:  
 Pred. No.: 1..96e-91 Length: 891  
 Score: 1012.50 Matches: 200  
 Percent Similarity: 82.53% Conservative: 22  
 Best Local Similarity: 74.35% Mismatches: 30  
 Query Match: 35.45% Indels: 17  
 DB: 14 Gaps: 5

US-10-019-783-2 (1-551) x CD437584 (1-891)

QY 106 HisAlaGluAspSerAlaAlaGluSerAsnGlyHisAlaAlaAlaGlu 125  
 Db 132 CACGCCAGCGAC-----AATGGC-----ACGGCGCAGCG----- 161

QY 126 GluGluGluAlaValGluTrpAsnGluAla-----GlyAlaLysAspGlyValLeuAla 144  
 Db 162 -----GAGTGGCGGTTCGCTCGGGCGGCAAGAGGCGCTCGCGCTG 206

QY 145 ThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGluGlu 164  
 Db 207 GCGGGGACAGATGATCGCGCGCGCGCGGTTCAGATCAGCGAGCGGTGACGGG 266

QY 165 LysGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 184  
 Db 267 CGGAGCCCGCGCGCGGTTCGCGGTTCGCGACGGGACCGCTCGGTGTTCCCGCGCTTC 326

QY 185 ArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsn 204  
 Db 327 CGCACCGCGCGGAGCGGAGGACCGGTGGCGCGCGCTCGGCGGCAAGTTCAAC 386

QY 205 CysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSer 224  
 Db 387 TCTTACCCCGCGCGGTGGCTCTCCCGAAGCCCGCGTGTCTGGCGGAGCACTGTCA 446

QY 225 GlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGln 244  
 Db 447 AGTGATCTTCCATACAAGGTGTCAACTGATGACATCTTCTCACCGCTGGAGGCACTCA 506

QY 245 AlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPro 264  
 Db 507 GCATTGAGGTGTGCTGTCTCAGTCTTGTCCCAA---CCGGCGCCCAACATTATTGCTCCCA 563

QY 265 ArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHis 284  
 Db 564 AGACAGCGCTATCCAAATTATGAGCGCGCTGCAGGACTGCACAACTTACAAGTTGCGCAT 623

QY 285 PheAspLeuIleProAspLysGlyTrpClnIleAspIleAspSerLeuGluSerIleAla 304  
 Db 624 TTCGATCTGATTCCTCGAGAGAGGGTGGAAATTCATATCGACTCTCTGGAGTCAATTGCT 683

QY 305 AspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyr 324  
 Db 684 GACAAAGAACACCCAGCGCAATGGTTCATCAATAACCCCAACATCTTCGCGCAGTGTCTAC 743

QY 325 SerTyrAspHisLeuSerLysValAlaGluValAlaLysArgIleLeuValIle 344  
 Db 744 ACCGTGAACATTATAGCCAGGTGCGGAGTACGAAGAGCTTGAATACTAGTATC 803

QY 345 AlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVal 364  
 Db 804 GCTGATGAAGTGTATGAAACCTGGTGTTCGGACACACCTTACGTCCTCCCAATGGGTGC 863

QY 365 PheGlyHisIleThrProValLeuSer 373  
 Db 864 TTGGCCATATTGCCCTGTGTGAGC 890

RESULT 7  
 LOCUS BQ466157 606 bp mRNA linear EST 30-MAY-2002  
 DEFINITION HT01L21T HT Hordeum vulgare subsp. vulgare cDNA clone HT01L21  
 5-PRIME, mRNA sequence.  
 ACCESSION BQ466157 GI:21273939  
 VERSION BQ466157.1  
 KEYWORDS EST.  
 SOURCE Hordeum vulgare subsp. vulgare  
 ORGANISM Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae  
 ; Triticeae; Hordeum.  
 1 (bases 1 to 606)  
 Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner  
 A.  
 TITLE Barley ESTs from germinating seeds  
 JOURNAL Unpublished  
 COMMENT Contact: Stein Nils  
 Molecular Markers Group, Department Genbank  
 Institute of Plant Genetics and Crop Plant Research (IPK)  
 Corrensstr. 3, 06466, Gatersleben, Germany  
 Tel: 039482-5592  
 Fax: 039482-5595  
 Email: stein@ipk-gatersleben.de  
 Insert Length: 606 Std Error: 0.00  
 Plate: 1 row: L column: 21  
 Seq primer: T3.  
 FEATURES  
 Location/Qualifiers  
 1..606  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="barke"  
 /db\_xref="taxon:112509"  
 /clone="HT01L21"  
 /tissue\_type="endosperm early"

```

/dev stage="0-16 hours after imbibition"
/lab_host="XL10-Gold"
/clone_lib="HT"
/notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present, as well as the EcoRI adapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
due to the cloning system used Blue/white selection for
recombinants is not 100% reliable."
BASE COUNT 113 a 216 c 190 g 87 t
ORIGIN

Alignment Scores:
Pred. No.: 4,95e-91 Length: 606
Score: 1006.00 Matches: 197
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.49% Mismatches: 0
Query Match: 35.22% Indels: 0
DB: 13 Gaps: 0

US-10-019-783-2 (1-551) x BQ466157 (1-606)
QY 150 SerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArgPro 169
|||
Db 11 AGCATCCGGGCGGATACGGTACAGATCAGCGGACGGTGCAGGAGAGGGCGCGGCC 70
|||
QY 170 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 189
|||
Db 71 GTGCTGCCGCTGGCCAGCGGACCGCTCCGTTTCCCGGGCTTCCGACGCGCGTCGAG 130
|||
QY 190 AlaGluAspAlaValAlaAlaValAlaArgThrGlyGlnPheAsnGlyTyrProAlaGly 209
|||
Db 131 GCGAGGACGCGTCCGCGCGCGTGGCGACCGGCTCACTGCTACCTACCGCGCGCG 190
|||
QY 210 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 229
|||
Db 191 GTCCGCTCCCGCGCGGACGAGCGCGTGGCGAGCACCTGTGCGAGGGCGTGGCGTAC 250
|||
QY 230 MetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValle 249
|||
Db 251 ATGCTATCGCGCGACGCTTCTCCACCGCGGACCCAGCGGATCGAGTACATA 310
|||
QY 250 IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 269
|||
Db 311 ATCCGGTGTGGCCAGACCGCGCGGCGCAACATTTCTGCTCCCGAGCGGCTACCCA 370
|||
QY 270 AsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIlePro 289
|||
Db 371 AACTACGAGCGCGCGCGCTTCAACAGGCTGGAGTCCGGCATTTTCGACCTCATCCCC 430
|||
QY 290 AspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 309
|||
Db 431 GACAAGGGTGGGAGATCGACATCGCTCGTGAATCCATCCGCGGACAGACACACCC 490
|||
QY 310 AlaMetValIleIleAsnProAsnAspProCysGlySerValTyrSerTyrAspHisLeu 329
|||
Db 491 GCCATGGTGCATATAAACCACCAACCGGTGGCGGACGCTTACTCTCAGCACCATCTG 550
|||
QY 330 SerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGlu 347
|||
Db 551 TCCAGGTGGGAGGTGGCGAAGGCTGGGATATTTGTTGATTGCTGACGAG 604
|||

RESULT 8
CB668396/c
LOCUS
DEFINITION OSJNE16D17.1 r OSJNE16D17.1 (japonica cultivar-group) cDNA
ACCESSION CB668396
VERSION OSJNE16D17.1
KEYWORDS EST.
SOURCE
Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 843)
Jantauriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 16 row: D column: 17
Seq primer: gga aac agc tat gac cat g.
FEATURES
source
1..843
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNE16D17"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNE16D17"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT 257 a 190 c 140 g 256 t
ORIGIN

Alignment Scores:
Pred. No.: 3,28e-90 Length: 843
Score: 1000.00 Matches: 186
Percent Similarity: 89.41% Conservative: 25
Best Local Similarity: 78.81% Mismatches: 25
Query Match: 35.01% Indels: 0
DB: 14 Gaps: 0

US-10-019-783-2 (1-551) x CB668396 (1-843)
QY 311 MetValIleIleAsnProAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330
|||
Db 843 ATAGTCATCATTAATCCCAATATCCATCCGCGGATGTGTACCTTACGAGCATTTATCC 784
|||
QY 331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350
|||
Db 783 AAGGTGCGAGAGTAGCAAGGAAGCTTGGCATATTGGTAATTACTGATGAGGTGATGCT 724
|||
QY 351 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370
|||
Db 723 AATTTGGTTTTGGGAGTTCCTTCCCAATGGGTGCTTTGGGCGACATCGTACCA 664
|||
QY 371 ValLeuSerIleGlySerLeuSerLysSerTyrIleValProGlyTyrArgLeuGlyTyr 390
|||
Db 663 ATATTAAACCATAGGATCGCTATCAAGAGGTGGATGTCGCGGATGCGGACTTGGTGG 604
|||
QY 391 ValAlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleThr 410
|||
Db 603 GTAGCAATATGTGACCCCAAGAGACTTACAAAGAAACCAAGATTGCAACATTAATTACT 544
|||
QY 411 AsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIle 430
|||
Db 543 AATTTCCCTTAATGTTTCACTGATCCAGCACTTTCATTCAGGAGGCTTACCGAATATT 484
|||

```



Kudrna,D.; Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea

TITLE  
JOURNAL  
COMMENT

Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: gta aaa cga cgg cca gtg  
BACKWARD: gga aac agc tat gac cat g  
Plate: 01 row: J column: 23  
Seq primer: gga aac agc tat gac cat g.

Location/Qualifiers  
1..902  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="IR36"  
/db\_xref="taxon:39945"  
/db\_xref="OSIIEA01U23"  
/tissue\_type="leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSIIEa"  
/note="Vector:pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
XhoI; Lesion Mimic SPL 11"

BASE COUNT 270 a 202 c 153 g 277 t  
ORIGIN

Alignment Scores:

Pred. No.:	3..63e-90	Length:	902
Score:	1000.00	Matches:	186
Percent Similarity:	79.41%	Conservative:	25
Best Local Similarity:	88.41%	Mismatches:	25
Query Watch:	35.01%	Indels:	0
DB:	14	Gaps:	0

US-10-019-783-2 (1-551) x CB618033 (1-902)

QY	311	MetVallelleAsnProAsnAsnPrOcysGlySerValTySertyrAspHisLeuSer	330
Db <td>901</td> <td>ATAGTCATCATAAATCCCAATAATCCATCGGGAAATGTACACTTACGACATTATACC</td> <td>842</td>	901	ATAGTCATCATAAATCCCAATAATCCATCGGGAAATGTACACTTACGACATTATACC	842
QY <td>331</td> <td>LysValalaGluValAlaLysArGleUGlylleLeuVallieAlaAspGluValTyrcGly</td> <td>350</td>	331	LysValalaGluValAlaLysArGleUGlylleLeuVallieAlaAspGluValTyrcGly	350
Db <td>841</td> <td>AAGTGCGCAGAGTAGCAGGAAGCTTGGGATATTGGTAATTACTGATGAGGTATGGT</td> <td>782</td>	841	AAGTGCGCAGAGTAGCAGGAAGCTTGGGATATTGGTAATTACTGATGAGGTATGGT	782
QY <td>351</td> <td>LysLeuValLeuGlySerAlaProPheilleProMetGlyValPheGlyHisIleThrPro</td> <td>370</td>	351	LysLeuValLeuGlySerAlaProPheilleProMetGlyValPheGlyHisIleThrPro	370
Db <td>781</td> <td>AATTGGTTTTTTGGGAGTTCGCCATTGTCCCAATGGGTTCCTTTGGGCACATCGTACCA</td> <td>722</td>	781	AATTGGTTTTTTGGGAGTTCGCCATTGTCCCAATGGGTTCCTTTGGGCACATCGTACCA	722
QY <td>371</td> <td>ValLeuSerlleGlySerLeuSerLySerTrpileValProGlylTrpArgLeuGlyTrp</td> <td>390</td>	371	ValLeuSerlleGlySerLeuSerLySerTrpileValProGlylTrpArgLeuGlyTrp	390
Db <td>721</td> <td>ATAATACCATAGATCGCTATCAAAGAGTGGATAGTCGGGATCGGATCGGACTTGGTGG</td> <td>662</td>	721	ATAATACCATAGATCGCTATCAAAGAGTGGATAGTCGGGATCGGATCGGACTTGGTGG	662
QY <td>391</td> <td>ValAlaValTyRAspProArGLySileleUGlnGluThrLysIleSerThrSerIleThr</td> <td>410</td>	391	ValAlaValTyRAspProArGLySileleUGlnGluThrLysIleSerThrSerIleThr	410
Db <td>661</td> <td>GTAGCAATATGTGACCCCAAGAAGACTCTACAAGAACCAAGATTGGCAACATTAAITACT</td> <td>602</td>	661	GTAGCAATATGTGACCCCAAGAAGACTCTACAAGAACCAAGATTGGCAACATTAAITACT	602
QY <td>411</td> <td>AsnTyrlEuasnValSerThrAspProAlaThrPhelleGlnlaAlaLeuProGlnile</td> <td>430</td>	411	AsnTyrlEuasnValSerThrAspProAlaThrPhelleGlnlaAlaLeuProGlnile	430
Db <td>601</td> <td>AATTTCCTTAATGTTTCACTATGATCCAGCAACTTTCATTTCAGGAGCTCTACCGAAATAT</td> <td>542</td>	601	AATTTCCTTAATGTTTCACTATGATCCAGCAACTTTCATTTCAGGAGCTCTACCGAAATAT	542
QY <td>431</td> <td>LeuGluAsnThrLySGluAspPhePheLySaLatlleiledGlyLeuLeuLysGluSerSer</td> <td>450</td>	431	LeuGluAsnThrLySGluAspPhePheLySaLatlleiledGlyLeuLeuLysGluSerSer	450
Db <td>541</td> <td>CTTAAGATACCAAGGAGAGATCTTTTAAGAGGNTAATTCATTTCCTTACGGAACATCA</td> <td>482</td>	541	CTTAAGATACCAAGGAGAGATCTTTTAAGAGGNTAATTCATTTCCTTACGGAACATCA	482
QY <td>451</td> <td>GluilleCysTyrlYsGlnlleLysGluAsnLystyrlleThrCysProHislYsProGlu</td> <td>470</td>	451	GluilleCysTyrlYsGlnlleLysGluAsnLystyrlleThrCysProHislYsProGlu	470

Db 491 GATATTTCCTTAGAGGAATAAAGATATTAAATGCATCACTTCCTCACAGCCCGAA 421

Qy 471 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluLeuAspAsp 490

Db 421 GGATCCATGTTTGTGATGGTGAATTTGAACCTATATCTTTTGGAGGGATCCATGATGAT 362

Qy 491 IleAspPheCysCysLysLeuAlaIysGlnGluSerValIleLeuCysProGlySerVal 510

Db 361 GTTATGATTTTGTTCACACTTGCAGAAAGAGTGGTGATTTCTTTGGCCAGGGAGTGTG 302

Qy 511 LeuGlyWetAlaAsnTpxValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 530

Db 301 CTGGGAATGAGAGATTGGTTCGCTACTTTTTCATTGATTCACTTCTCTCCTGGAT 242

Qy 531 GlyLeuGlyArgIleLysSerPheCysGlnArgAsnLysLysArgAsn 546

Db 241 GGTCTTGAGAGGATCAAAATCCTTCTGCCAAAGGCACAGAGAAAAAC 194

RESULT 11

CB6711005/c

LOCUS

DEFINITION CB6711005 849 bp mRNA linear EST 09-APR-2003

clone OSUNEE04F21.3', mRNA sequence.

ACCESSION CB6711005

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE

AUTHORS 1 (bases 1 to 849)

Jantasuriyarat C., Lu G., Gowda M., Hatfield J., Zhou B., Mazur E., Kudrna D., Dean R., Soderlund C., Wang R. and Wang G.

TITLE Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea

JOURNAL Unpublished

COMMENT Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 488A, P.O. Box 210088, Tucson, AZ 85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aaa cga cgg cca gtg

BACKWARD: gga aac agc tat gac cat g

Plate: 04 row: F column: 21

Seq primer: gga aac agc tat gac cat g.

FEATURES

source

Location/Qualifiers

1..849

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="OSUNEE04F21"

/tissue\_type="Leaf"

/dev\_stage="3 week"

/lab\_host="DH10B"

/clone\_lib="OSUNEE"

/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2: XhoI; 24 hrs after immunoculation with Rice Blast (70-15)"

BASE COUNT 260 a 190 c 140 g 259 t

ORIGIN

Alignment Scores:

Pred. No.: 4.18e-90 Length: 849

Score: 993.00 Matches: 186

Percent Similarity: 89.36% Conservativeness: 24

Best Local Similarity: 79.15% Mismatches: 25

Query Match: 34.98% Indels: 0

DB: 14 Gaps: 0

US-10-019-783-2 (1-551) x CB671005 (1-849)

QY 312 ValIleLeuAsnProAsnAspProCysGlySerValTyrSerTyrAspHisLeuSerLys 331.  
 Db 847 GTCATCATTAATCCCAATATCCATCCATCGGGAATGTATACATTCACGAGCATTTATCCAAG 788  
 QY 332 ValAlaGluValAlaLysArgLeuGlyLysLeuValIleLeuValIleAlaAspGluValTyrGlyLys 351  
 Db 787 GTGGCAGAGGTAGCAAGAACTTGGGATATTGGTAATTCTGATGAGGTGTTGTTGTAAT 728  
 QY 352 LeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrProVal 371  
 Db 727 TTGGTTTGGGAGTCCCATTTGTCCTAATGGTTGCTTGGGCACATCGTACCAATA 668  
 QY 372 LeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpVal 391  
 Db 667 TTAACCATAGGATCGCTATCAAGAGGTGGATGTGCGGATGGCCACTTGGTTGGGTA 608  
 QY 392 AlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsn 411  
 Db 607 GCATATGTGACCCCAAGAACTCTACAGAAACCAAGATTGCAACATTAATTACTAAT 548  
 QY 412 TyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeu 431  
 Db 547 TTCCTTAATGTTCAACTGTATCCAGCACTTTCATTCCGGGAGCTCTACCGAATATTCTT 488  
 QY 432 GluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGlu 451  
 Db 487 AAGAAATACCAAGGAAGATTCTTTAAGGAGTAATATTGATTGCTTACGGAACATCATGAT 428  
 QY 452 IleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGly 471  
 Db 427 ATTTGCTATAGAGGAATAAAGCATATTAAATGCATCACTTGTCTCAAGCCGCAAGGA 368  
 QY 472 SerMetPheValMetValLysLeuAsnLeuHisLeuGluGluIleAspAspAspIle 491  
 Db 367 TCCATGTTTGGTGGTGAATTTGAACCTATATCTTTTGGAGGAATCCATGATGATT 308  
 QY 492 AspPheCysLysLeuAlaLysGluLysSerValIleLeuCysProGlySerValLeu 511  
 Db 307 GATTTTGTGCCAACTTCGGAAGAGAGAGTCGTGATCTTTGCCAGGAGGTGCTG 248  
 QY 512 GlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGly 531  
 Db 247 GGAATCAAGAAATGGGTTTCGATTACTTTTCTATTGATTCTCTCTCTGATGTT 186  
 QY 532 LeuGlyArgIleLysSerPheCysGlnArgAsnLysLysArgAsn 546  
 Db 187 CTTGAGAGATCAATCTCTTTCTGCCAAGGACACAGAGAAAC 143

RESULT 12  
 CB618032  
 LOCUS  
 DEFINITION OSIIIEa01J23.f OSIIIEa Oryza sativa (indica cultivar-group) cDNA  
 clone OSIIIEa01J23 5', mRNA sequence.  
 ACCESSION CB618032  
 VERSION CB618032.1 GI:29613019  
 KEYWORDS EST.  
 SOURCE Oryza sativa (indica cultivar-group)  
 ORGANISM Oryza sativa (indica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzoae; Oryza.  
 REFERENCE 1 (bases 1 to 809)  
 AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
 Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
 TITLE Large-scale identification of ESTs involved in the interaction  
 between rice and Magnaporthe grisea  
 JOURNAL Unpublished  
 COMMENT Contact: Rod Wing  
 Arizona Genomics Institute  
 University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
 85721-0088, USA  
 Tel: 520 626 3967  
 Fax: 520 621 9288  
 Email: http://genome.arizona.edu

PCR Primers  
 FORWARD: gta aaa cga cgg cca gtc  
 BACKWARD: gga aac agc tat gac cat g  
 Plate: 01 row: J column: 23  
 Seq primer: gta aaa cga cgg cca gtc.

#### FEATURES

source location/Qualifiers

1..809  
 /organism="Oryza sativa (indica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="IR36"  
 /db\_xref="taxon:39946"  
 /clone="OSIIIEa01J23"  
 /tissue\_type="Leaf"  
 /dev\_stage="3 week"  
 /lab\_host="DH10B"  
 /clone\_lib="OSIIIEa"  
 /note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2:  
 XhoI; Lesion Mimic SPL II"  
 BASE COUNT 176 a 245 c 235 g 153 t  
 ORIGIN

#### Alignment Scores:

Pred. No.: 6,948-90 Length: 809  
 Score: 996.50 Matches: 200  
 Percent Similarity: 76.08% Conservative: 29  
 Best Local Similarity: 66.45% Mismatches: 36  
 Query Match: 34.89% Indels: 36  
 DB: 14 Gaps: 5

US-10-019-783-2 (1-551) x CB618032 (1-809)

QY 67 HisAlaGluAlaThrAlaAsnGlyHisGly-GluAlaThrAlaAsnGlyLysThrAsnG1 86  
 Db 9 CACCAACACAAAACAGACACGCCATGGCACCGACGCGG----- 51  
 QY 86 yHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluHi 106  
 Db 52 -----GCGCGCGCGCGACGACGCAACGCG----- 75  
 QY 106 salAGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaAlaGluCl 126  
 Db 76 -----GCGCGCGAGACGCGC----- 93  
 QY 126 uGluGluAlaValGluTrpAsnPhe---AlaGlyAlaLysAspGlyValLeuAlaAlaTh 145  
 Db 94 ----AGCAGCAAGGAGTGGAGGCTGACGCGCGCGACGAGGCGCGGATGGCGCGCGC 149  
 QY 145 rGlyAlaAsnMetSerIleArgAlaIleArgTyrIleSerAlaSerValGlnGluLy 165  
 Db 150 GGGGGACAAGATGAGCATCCGGCGCGGTGGGTACAGATCAGCGCAGGTCGACGACCG 209  
 QY 165 sGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheAr 185  
 Db 210 CGGCGCGCGCGCGCTCTCGCGCTCGCCACGCGACCCCTCCGTGTTCCCGAGTTCCG 269  
 QY 185 gThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsnCy 205  
 Db 270 CACCGCGCGCGCGAGGAGCGCGTCCCGACCGCTCCGCTCCGCGACCTTCAACTG 329  
 QY 205 sTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerG1 225  
 Db 330 CTRACCCCGCGCGCTCGGCTCTCCCGCGCGCGAGCGTGTGTGGCAGATCACTTTCAGC 389  
 QY 225 nGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAl 245  
 Db 390 CGACCTCCCATACAGCTATCTTCTGATGACATCTTCTAACCGCTGGAGGAACCTCAGC 449  
 QY 245 alleGluValIleIleProValAlaGlnThrAlaGlyAlaAsnIleLeuLeuProAr 265



```

Db      450 CATCGAGGTCGTAATCTCAATCTTGCCTTGCCTTCCCTAG 506
QY      265 gProGlyTyrProAsnTyrGluAlaAargAlaPheAsnArgLeuGluValArgHisPh 285
Db      507 ACCAGGCTACCCAACTATGAGCTCGAGCGGTTCAACAACCTTGAAGTTCGTCATT 566
QY      285 eAspLeuileProAspLysGlyTyrGluLeuAspLeuSerLeuAlaAs 305
Db      567 TGATCTTATCTCGAAGGCTGGAGATTGACCTTAACCTCCCTAGTAATTAATCGGA 626
QY      305 pYsAsnThrAlaMetValleleAsnProAsnAsnProCysGlySerValTyrSe 325
Db      627 CAAGAACACTACTGCGCAATGATCATATAATCCCAATATCCATCGCGGAATGCTACAC 686
QY      325 rTyAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyLeuValleleAl 345
Db      687 TTACAGCAATTTATCCAGGTGGCAGAGTACAGGAAGCTTGGGATATTGGTAATTAC 746
QY      345 aAspGluValTyrGlyLysLeuValleuGlySerAlaProPheileProMetGlyValph 365
Db      747 TGATGAGGTGATGTGTAATTTGGTTTGGGAGTTCCCAATTTGCCCCAATGGGTGCTT 806
QY      365 e 365
Db      807 T 807

RESULT 13
CB652840/c
LOCUS
DEFINITION
OSUNEC03E07.r OSUNEC Oryza sativa (japonica cultivar-group) cDNA
clone OSUNEC03E07 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 832)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga asc agc tat gac cat g
Plate: 03 row: E column: 07
Seq primer: gga aac agc tat gac cat g.
Location/Qualifiers
1. 832
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNEC03E07"
/tissue_type="Leaf"
/dev stage="3 week"
/lab host="DH10B"
/clone_lib="OSUNEC"
note=Vector; pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
256 a 186 c 140 g 250 t
BASE COUNT
source

```

## ORIGIN

## Alignment Scores:

Pred. No.: 1.03e-89 Length: 832  
 Score: 995.00 Matches: 185  
 Percent Similarity: 89.32% Conservative: 24  
 Best Local Similarity: 79.06% Mismatches: 25  
 Query Match: 34.84% Indels: 0  
 DB: 14 Gaps: 0

US-10-019-783-2 (1-551) x CB652840 (1-832)

```

QY      313 lIleleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysVal 332
Db      831 ATATAAATCCCAATATGAGCTCGAGGAATGTGTACATTTACGAGCATTTATCCAAAGTG 772
QY      333 AlaGluValAlaLysArgLeuGlyLleLeuValleleAlaAspGluValTyrGlyLysLeu 352
Db      771 GCAGAGGTAGCAAGGAAGCTTGGGATATTGGTAATTACTGATGAGGTGTATGGTAATTG 712
QY      353 ValLeuGlySerAlaProPheileProMetGlyValPheGlyHisIleThrProValLeu 372
Db      711 GTTTTGGGAGTTCCCAATTTGTCCTCAATGGGTGCTTTGGGCACATCGTACCAATATTA 652
QY      373 SerIleGlySerLeuSerLysSerTyrPheValProGlyTyrArgLeuGlyTyrValAla 392
Db      651 ACCATAGGATCGCTATCAAGAGGTGTAGTAGTCCCGGATGGGCACTTGGTGGGTAGCA 592
QY      393 ValTyrAspProArgLysIleLeuGluThrLysIleSerThrSerIleThrAsnTyr 412
Db      591 ATATGTGACCCCAAGAAGACTCTACAGAAACCAAGATTGCAACATTAATTACTAATTTTC 532
QY      413 LeuAsnValSerThrAspProAlaThrPheileGluAlaLeuProGluIleLeuGlu 432
Db      531 CTTAATGTTTCAACTGATCCAGCACTTTCATTACGGGAGCTCTACCGAATATTTCTTAAG 472
QY      433 AsnThrLysGluAspPhePheLysAlaIleleGlyLeuLysGluSerSerGluIle 452
Db      471 AATACCAAGGAAGAAATTCITTAAGAGGATAATGATTGCTTCTACGAAACATCAGATATT 412
QY      453 CysTyrLysGluLysGluAsnLysTyrIleThrCysProHisLysProGluGlySer 472
Db      411 TGTATAGAGGAATAAAGGATATTAAATGATCATCACTTGTCTCACAAGCCCAAGGATCC 352
QY      473 MetPheValMetValLysLeuAsnLeuHisLeuGluGluLeuAspAspAspIleAsp 492
Db      351 ATGTTTGTGATGTGAATTTGAACCTATATCTTTTGGAGGGAATCCATGATGTTGAT 292
QY      493 PheCysLysLeuAlaLysGluSerValleleLeuCysProGlySerValLeuGly 512
Db      291 TTTTGTGCAACTTCGGAAGAGAGTCCGTGATTTTTCGCCAGGAGTGTGCTGGGA 232
QY      513 MetAlaAsnTyrValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeu 532
Db      231 ATGAAGAATTGGTTCGCATTACTTTTGTATGATGATCACTTCTCTCTCTGATGCTTT 172
QY      533 GlyArgIleLysSerPheCysGlnArgAsnLysLysArgAsn 546
Db      171 GAGAGGATCAATCTTCTGCCAAGAGGCACAGAAGAAAAC 130

```

## RESULT 14

## CB646007/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Oryza sativa (japonica cultivar-group)

## Oryza sativa (japonica cultivar-group)

## Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

## Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

## Ehrhartoideae; Oryzeae; Oryza.

## Ehrhartoideae; Oryzeae; Oryza.

## Ehrhartoideae; Oryzeae; Oryza.

## Ehrhartoideae; Oryzeae; Oryza.

## Ehrhartoideae; Oryzeae; Oryza.

```

REFERENCE
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
JOURNAL between rice and Magnaporthe grisea
COMMENT Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 08 row: C column: 06
Seq primer: gga aac agc tat gac cat g.
FEATURES
Location/Qualifiers
source
1..883
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEB08C06"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEB"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (Che 86061
)"
BASE COUNT 265 a 199 c 151 g 268 t
ORIGIN
Alignment Scores:
Pred. No.: 2,89e-88 Length: 883
Score: 981.00 Matches: 182
Percent Similarity: 88.84% Conservative: 25
Best Local Similarity: 78.11% Mismatches: 26
Query Match: 24.35% Indels: 0
DB: 14 Gaps: 0
US-10-019-783-2 (1-551) x CB646007 (1-883)
QY 314 IleAsnProAsnAspProCysGlySerValTyrSerTyrAspHisLeuSerLysValAla 333
Db 861 ATAAATCCCAATATCCATGCGGNAATGTACACTTACGAGCATTTATCCAGGTGGCA 822
QY 334 GluValAlaLysArgLeuGlyLeuValLeuValLeuValLeuValTyrGlyLysVal 353
Db 821 GAGGTAGCAAGAACCTGGGATATGTTAATTAAGTATGAGGTATGTTAATTTGGTT 762
QY 354 LeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSer 373
Db 761 TTTGGAGTCCCAATTTGTCCTCAATGGTGTCTTTGGCCACATCCGACCAATTAACC 702
QY 374 IleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaVal 393
Db 701 ATAGGATCGCTATCAAGAGGTGGATAGTGGCGGATGGCGACTGGGTGGGTAGCAATA 642
QY 394 TyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeu 413
Db 641 TGTGACCCCAAGAACTCTACAGAAACCAAGATTGCAATTAATTAATTAATTTCTT 582
QY 414 AsnValSerThrAspAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsn 433
Db 581 ATGTTTCACTGATCCAGCACTTTTCATTCAGGGAGCTCTACCGAATATTCTTAAGAA 522
QY 434 ThrLysGluAspPhePheLysAlaIleLeuGlyLeuLeuLysGluSerSerGluIleCys 453
Db 521 ACCAAGGAAGAATCTTTTAAGAGGATAATTGTTCTTCGGAACATCATGATATTTCG 462

```

```

QY 454 TyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMet 473
Db 461 TATAGAGAAATAAGGATATTAAATGATCATCTGTCTCTACAGCCCGAAGGATCCATG 402
QY 474 PheValMetValLysLeuAsnLeuHisLeuLeuGluGluLeuAspAspAspIleAspPhe 493
Db 401 TTTGTGATGGTCAAAATTCGAACATATCTTTTGGAGGAATCCATGATGTTGATTTT 342
QY 494 CysCysLysLeuAlaLysGluLeuValIleLeuCysProGlySerValLeuGlyMet 513
Db 341 TGTTCGCCAACTTCGAAAGAGTGGGTGATTCTTTGCCAGGAGTGTCTCTGGGATG 282
QY 514 AlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGly 533
Db 281 AAGATTGGGTTCGCATTACTTTTCTATTGATTCACTTCTCTCTGGATGTTCTTGG 222
QY 534 ArgIleLysSerPheCysGlnArgAsnLysLysArgAsn 546
Db 221 AGGATCAAAATCTCTCTGCAAGGCAAGAAAGAAAC 183
RESULT 15
CB670913 792 bp mRNA linear EST 09-APR-2003
LOCUS OSJNEB04D20.f OSJNE Oryza sativa (japonica cultivar-group) cDNA
DEFINITION clone OSJNEB04D20 5', mRNA sequence.
ACCESSION CB670913
VERSION CB670913.1 GI:29674638
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE 1 (bases 1 to 792)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
JOURNAL between rice and Magnaporthe grisea
COMMENT Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 04 row: D column: 20
Seq primer: gta aaa cga cgg cca gtc.
FEATURES
Location/Qualifiers
source
1..792
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEB04D20"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEB"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"
BASE COUNT 177 a 242 c 229 g 144 t
ORIGIN
Alignment Scores:
Pred. No.: 7.08e-87 Length: 792
Score: 966.50 Matches: 195
Percent Similarity: 76.19% Conservative: 29

```

Best Local Similarity: 66.33% Mismatches: 34  
 Query Match: 33.84% Indels: 36  
 DB: 14 Gaps: 5

US-10-019-783-2 (1-551) x CB670913 (1-792)

```

QY 67 HisAlaGluAlaThrAlaAsnGlyHisGly-GluAlaThrAlaAsnGlyLysThrAsnG1 86
   |||
   |||
   |||
Db 15 CACCACAAACACAGACGACGCGCATGACCGACGACGCGG----- 57
   |||
QY 86 yHisargGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluH1 106
   |||
   |||
   |||
Db 58 -----GCGGCGGCGGAGACGACCAACGCG----- 81
   |||
QY 106 sAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGluG1 126
   |||
   |||
   |||
Db 82 -----GCGGCGGAGAGCGACGCG----- 99
   |||
QY 126 uGluGluAlaValGluTrpAsnPhe---AlaGlyAlaLysAspGlyValLeuAlaAlaTh 145
   |||
   |||
   |||
Db 100 ----AGCAGCAAGAGTGAGGCTGACGCGCGCGACGAGGCGCGCGCATGGCGGCGGC 155
   |||
QY 145 rGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLyl 165
   |||
   |||
   |||
Db 156 GGGGACAAAGATGAGCATCCGGCGGTCGGTACAGATCAGCGCGCGCTCGCGGACTTCAACTG 215
   |||
QY 165 sGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheAr 185
   |||
   |||
   |||
Db 216 CGGCGCGCGCGCGCTCGCGCTCGCCACGCGCGACCCCTCCGTCGTCGCGAGTTCGG 275
   |||
QY 185 gThrAlaValGluAlaGluAspAlaValAlaAlaValAlaValArgThrGlyGlnPheAsnCy 205
   |||
   |||
   |||
Db 276 CACCGCGCGCGAGGCGGAGGACGCGCTCGCGACGCGCTCGCGGACTTCAACTG 335
   |||
QY 205 sTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerG1 225
   |||
   |||
   |||
Db 336 CTACCGCGCGCGCTCGCGCTCCCGCGCGCGAGCTGCTGGCGAGATCATTTGTACG 395
   |||
QY 225 nGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAl 245
   |||
   |||
   |||
Db 396 CGACCTCCCATACAAAGTATCTTCGTGATGATCTTCCTAACCGCTGGAGGAACCTCAGGC 455
   |||
QY 245 aileGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProAr 265
   |||
   |||
   |||
Db 456 CATCGAGGTGTAATCTCAATCTTGGCCAA---CCTGGCACAAACATATTGCTTCCTAG 512
   |||
QY 265 gProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValAlaArgHisph 285
   |||
   |||
   |||
Db 513 ACCAGGCTACCCAAACTATGAAGCTCGAGCGCGGTTCAACAACCTTGAGTTCGTCACCT 572
   |||
QY 285 eAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAs 305
   |||
   |||
   |||
Db 573 TGATCTTATTCCTGAGAAGGCTGGGAGATTGACCTTAACCTCCCTAGAAATCTATTGGCGGA 632
   |||
QY 305 pLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSe 325
   |||
   |||
   |||
Db 633 CAAGACACTACTGCGATAGTCATCATATATCCCAATATCCATGCGGGAATGTTACAC 692
   |||
QY 325 rTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAl 345
   |||
   |||
   |||
Db 693 TTACGAGCATTTATCCAAAGGTGGCAGAGGTAGCAAGGAAGCTTGGGATATTGGTAATTAC 752
   |||
QY 345 aAspGluValTyrGlyLysLeuValLeuGlySerAlaPro 358
   |||
   |||
   |||
Db 753 TGATGAGGTATGTAATTGGTTTGGGAGTTCCCA 792

```

Search completed: November 9, 2003, 01:20:06  
 Job time : 3372.72 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 23:39:05 ; Search time 488.386 Seconds  
(without alignments)  
3600.055 Million cell updates/sec

Title: US-10-019-783-2

Perfect score: 2856

Sequence: 1 MATVRSQGVAAVLAA.....LGRKSFQCKNKNSSDDC 551

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2141354 seqs, 1595478879 residues

Total number of hits satisfying chosen parameters: 4282708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USFIO.spool/US10019783/runat.07112003.161647.19988/app.query.fasta.1.1358  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10019783@cgn1.1.541@runat.07112003.161647.19988  
-NCPU=6 -ICPU=3 -NO\_XMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEVTIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/2/pubpna/PCN_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubpna/PCUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
17:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	1118.5	39.2	1389	10 US-09-938-842A-1984
				Sequence 1984, Ap

2	998	34.9	1338	10	US-09-938-842A-1989	Sequence 1869, Ap
3	663	23.2	858	9	US-09-770-445-608	Sequence 408, App
4	637.5	22.3	2712	11	US-09-919-039-379	Sequence 379, App
5	637.5	22.3	2754	14	US-09-880-107-3720	Sequence 3720, A
6	634.5	22.2	2921	14	US-10-198-846-11998	Sequence 11998, A
7	605.5	21.2	1631	14	US-10-198-846-11531	Sequence 11531, A
8	399	14.0	870	14	US-10-198-846-8896	Sequence 8896, A
9	359.5	12.6	1224	14	US-10-156-761-4891	Sequence 4891, Ap
10	359.5	12.6	9025608	14	US-10-156-761-4891	Sequence 1, Appli
11	348	12.2	1185	10	US-09-905-173-39	Sequence 39, Appli
12	341.5	12.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
13	339	11.9	370	10	US-09-878-574-2935	Sequence 2935, Ap
14	339	11.9	1257	10	US-09-738-626-3134	Sequence 3134, Ap
15	329	11.5	9025608	14	US-10-156-761-1	Sequence 1, Appli
16	313	11.0	1209	14	US-10-156-761-3272	Sequence 3272, Ap
17	311	10.9	1194	14	US-10-156-761-4501	Sequence 4501, Ap
18	307	10.7	837	14	US-10-198-846-12727	Sequence 12727, A
19	303.5	10.6	1206	14	US-10-156-761-1996	Sequence 1996, Ap
20	303.5	10.6	1247	10	US-09-971-536-12	Sequence 12, Appli
21	298.5	10.5	21252	10	US-09-070-927A-94	Sequence 94, Appli
22	289	10.1	1107	12	US-10-288-930-4	Sequence 4, Appli
23	289	10.1	1830121	14	US-10-329-960-1	Sequence 1, Appli
24	287.5	10.1	1122	10	US-09-905-173-18	Sequence 18, Appli
25	287.5	10.1	1122	14	US-10-060-432-18	Sequence 18, Appli
26	286	10.0	1170	9	US-09-967-645-2	Sequence 2, Appli
27	286	10.0	3614	10	US-09-070-927A-185	Sequence 185, App
28	283.5	9.9	1170	12	US-10-288-930-5	Sequence 5, Appli
29	283.5	9.9	1409	10	US-09-971-536-11	Sequence 11, Appli
30	271	9.5	1209	10	US-09-738-626-911	Sequence 911, App
31	258.5	9.1	1221	10	US-09-938-842A-1197	Sequence 1197, Ap
32	255.5	8.9	1197	10	US-09-905-173-21	Sequence 21, Appli
33	255.5	8.9	1197	14	US-10-060-432-21	Sequence 21, Appli
34	247	8.6	2127	10	US-09-908-928-1	Sequence 1, Appli
35	246	8.6	1572	10	US-09-908-928-3	Sequence 3, Appli
36	243	8.5	446	9	US-09-770-444-786	Sequence 786, App
37	242.5	8.5	870	14	US-10-198-846-2949	Sequence 2949, Ap
38	240.5	8.4	1076	10	US-09-070-927A-835	Sequence 835, App
39	236.5	8.3	764	10	US-09-974-300-232	Sequence 292, App
40	233.5	8.2	13884	10	US-09-070-927A-341	Sequence 341, App
41	231	8.1	730	9	US-09-770-149-110	Sequence 110, App
42	228.5	8.0	1495	11	US-09-746-660A-3	Sequence 3, Appli
43	226	7.9	458	9	US-09-770-444-496	Sequence 496, App
44	224.5	7.9	3309400	10	US-09-738-626-1	Sequence 1, Appli
45	224	7.8	1975	14	US-10-284-985-4	Sequence 4, Appli

## ALIGNMENTS

### RESULT 1

US-09-938-842A-1984  
; Sequence 1984, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1984  
; LENGTH: 1389  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana

```
US-09-938-842A-1984
Alignment Scores:
Pred. No.: 2,55e-110 Length: 1389
Score: 1118.50 Matches: 207
Percent Similarity: 65.94% Conservative: 97
Best Local Similarity: 44.90% Mismatches: 140
Query Match: 39.16% Indels: 17
DB: 10 Gaps: 3

US-10-019-783-2 (1-551) x US-09-938-842A-1984 (1-1389)
QY 89 GluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluHisAlaGlu 108
DB 7 GAAGAACAACACACACGCAATCTAGCGGTTCCCGCGTTTAAACTGAGAAAGAGCCCAT 66
QY 109 AspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGluGluGlu 128
DB 67 ACGCAAAACCAATGTCAGTAGC----- 93
QY 129 AlaValGluTrpAsnPheAlaGlyAlaAlaAspGlyValLeuAlaAlaThrGlyAlaAsn 148
DB 94 -----GTTTCGGCTTCGGTGAAGTAAAG-----GCAGCGAAGCATCCACC 138
QY 149 MetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArg 168
DB 139 GTAACGCTTAGAGTGTCATCTACATGCTCTTCGACAACTCGCGCAAGACGTCATTAAG 198
QY 169 ProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaVal 188
DB 199 ACCATTTTACCCTCGCGCACCGGTGACCTTCCTGCTACCTGCTTCCGTTACCTGATC 258
QY 199 GluAlaGluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsnCysTyrProAla 208
DB 259 GAAGCTGAAGACGGCGTCTCGAGTCTCGCTCGCGCAAGCAATCTTACGGTCCC 318
QY 209 GlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValPro 228
DB 319 GGAGCTGGGATCTCCCGCGCAAGACGCGCTTCTGCTATATATGAACCGAGATCTTCG 378
QY 229 TyrMetLeuSerAlaAspValPheLeuThrAlaGlyThrGlnAlaIleGluVal 248
DB 379 CACAAGTTAACCGCTGAGATATTTCTGACCGCTGATGCAACCAAGGATAGATC 438
QY 249 IleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyr 268
DB 439 GTTGCTGAATCGTTGGCTCGA-----CCAAACGCAACATCTTGCTCCACGCTCGGCTC 495
QY 269 ProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIle 288
DB 496 CCTCACTACGCGCTCGTGTCTTACAGTGTCTCGAGGTTCGCAAGTTTGATCTTCT 555
QY 289 ProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThr 308
DB 556 CCCGAGAAGAATGGAGATTCATCTTGAAGTATCGAGCCATTCAGACGAGAACAT 615
QY 309 ThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHis 328
DB 616 GTGGCTATGTTGTAATTAACCCCAACAATCCCTGTGGAAATGCTACTCTCAGCACCAT 675
QY 329 LeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluVal 348
DB 676 CTCAAAAAGGTTCAGACGCGCTAGGAAGTTCGGGATAATGGTGATCTCAGAGAAATA 735
QY 349 TyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIle 368
DB 736 TATGACCGAATCTATATTCGAGACAAATCCATTTGTTCAATGGGAAAGTTTGTTCGATA 795
QY 369 ThrProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeu 388
DB 796 GTCCCTGTATTGACACTAGCAGGATATCTAAGGATGGGTGTGTTCTCGATGGAATAAT 855
QY 389 GlyTyrValAlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSer 408

US-09-938-842A-1869
Sequence 1869, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krebs, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1869
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-1869

Alignment Scores:
Pred. No.: 2,29e-97 Length: 1338
Score: 999.00 Matches: 1394
Percent Similarity: 65.51% Conservative: 89
Best Local Similarity: 44.91% Mismatches: 139
Query Match: 34.94% Indels: 10
DB: 10 Gaps: 3

US-10-019-783-2 (1-551) x US-09-938-842A-1869 (1-1338)
QY 111 AlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGluGluGluAlaVal 130
```

Db 4 GCAGCAACGGAGTTACCACTGTAACGCAACGCAATGTT----- 45  
 Qy 131 GluTrpAspPheAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMetSer 150  
 Db 46 ---TGCGGTCAAAGCA-----AACGGTGCA---ACGAGTCATGCGCGCGGTGAGC 93  
 Qy 151 IleArgAlaIleArgTyrLysIleSerAlaSerValGlnGlnLysGlyProArgProVal 170  
 Db 94 TTGAGAAAGCTTCTTTGGGATGTTAAAACTGCACCATGAAACAGTGAAGAACCAAT 153  
 Qy 171 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 190  
 Db 154 TTGTTCCCAACTCCCGCGAGCGCTCCGCCATTCACACTTCAGGACTTGCCCGAAGCC 213  
 Qy 191 GluAspAlaValAlaAlaValArgThrGlyGlnPheAsnGlyValProAlaGlyVal 210  
 Db 214 GAGGAGCGCTTGCCGCACTGCACGCTCCGCGATGGCTAACTTTCACGCCACCGCCT 273  
 Qy 211 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMet 230  
 Db 274 GGAGTTTTCAGGCTAGAGGCGGTGGCTGAATATTTAAACGGAGAACTTCCGAGCAAG 333  
 Qy 231 LeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIlele 250  
 Db 334 CTGAAGCGCGAGATGTATATACCGGAGATGTAACCAAGCCATAGATCGTGATA 393  
 Qy 251 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 270  
 Db 394 GATTCCTTGCGGAATCCATCCGCCACATCTCTACTTCCAGCCGGGATATCTCTCAC 453  
 Qy 271 TyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProAsp 290  
 Db 454 TACGATGCTCGTGCTGTATAGCGGCTCGAGATTCGCAATACGATCTTCTCCCGAG 513  
 Qy 291 LysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrAla 310  
 Db 514 AGTGATTGGGAATCAATCTCATCGCTCGAGCGGCTCGGATGAGATACCGTCCGA 573  
 Qy 311 MetValIleLeuAsnProAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330  
 Db 574 ATGGTAATCATCAACCCCAACATCCATGTTGGAAACGCTACACCTACGACCATCTCAAC 633  
 Qy 331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350  
 Db 634 AAGTCGCGGAGATGGCTAGAAATCGTATATGATAATATCCGAGCAAGTATATGAT 693  
 Qy 351 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370  
 Db 694 CATGTTGTATATGAGACAGCCCTTATTCCTCCATGGGAAGTTTGCATCATAGCTCCG 753  
 Qy 371 ValLeuSerIleGlySerLeuSerIleValProGlyTyrArgLeuGlyTyr 390  
 Db 754 GTGATCACGCTCGGATCCATATCCAAAGATGGGTCAACCCAGGCTGGAGATTGGCTGG 813  
 Qy 391 ValAlaValTyrAspProArgLysIleLeuGlnGlnLysIleSerThrSerIleThr 410  
 Db 814 ATCCCATGACGATCCTTAAGTATCTTTGTATCTACAGGGGTAGTTCAGCAATAGAG 873  
 Qy 411 AsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIle 430  
 Db 874 GATTTCCTTGATTAACTCCACAGCTTCAATTTATTCCTCAGGAAGCACTTCTGTGATA 933  
 Qy 431 LeuGluAsnThrLysGluAspPhePheIleAlaIleIleGlyLeuLeuLysGluSerSer 450  
 Db 934 TTGAGAAAACACCTTAAGAGTTCTTCGAGAAGAAGATCAAAAGCCATGAGACGACGTC 993  
 Qy 451 GluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGlu 470  
 Db 994 GAGCTTCATGTGAGAGCTCAAGATATCTTGTCTCTTTGTCTCCCAAGAAACCCGAA 1053  
 Qy 471 GlySerMetPheValMetValLysLeuAsnHisLeuGluGluIleAspAsp 490

Db 1054 TCTTGTCTTATTATGTTGAGCTTGACACATCAATGTTGAATAATATCAAAATGAT 1113  
 Qy 491 IleAspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerVal 510  
 Db 1114 TTTGATTTCTGCAGCAAGCTAGTTAGTGAGAGAGTCTTATCTTATACGAGGAGTGGCT 1173  
 Qy 511 LeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 530  
 Db 1174 CTAGGGCAGAGAAATGGGTGAGGATATCGATAGGAACCGCAATCAGTGGTACAAGAA 1233  
 Qy 531 GlyLeuGlyArgIleLysSerPheCysGlnArgAsn 542  
 Db 1234 ATATTTCAGACACTAAAGGTTTCTATGATCGTCAT 1269

RESULT 3

US-09-770-445-608/c  
 ; Sequence 608, Application US/09770445  
 ; Patent NO. US2002023281A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gorlach, Jorn  
 ; APPLICANT: An, Yong-Qiang  
 ; APPLICANT: Hamilton, Carol M.  
 ; APPLICANT: Price, Jennifer L.  
 ; APPLICANT: Raines, Tracy M.  
 ; APPLICANT: Yu, Yang  
 ; APPLICANT: Rameaka, Joshua G.  
 ; APPLICANT: Page, Amy  
 ; APPLICANT: Matthew, Abraham V.  
 ; APPLICANT: Ledford, Brooke L.  
 ; APPLICANT: Woessner, Jeffrey P.  
 ; APPLICANT: Haas, William David  
 ; APPLICANT: Kricker, Maja  
 ; APPLICANT: Slader, Ted  
 ; APPLICANT: Davis, Keith R.  
 ; APPLICANT: Allen, Keith  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Hurlan, Patrick  
 ; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
 ; FILE REFERENCE: thaliana  
 ; CURRENT APPLICATION NUMBER: US/09/770,445  
 ; CURRENT FILING DATE: 2001-01-26  
 ; PRIOR APPLICATION NUMBER: US 60/178,472  
 ; PRIOR FILING DATE: 2000-01-27  
 ; NUMBER OF SEQ ID NOS: 999  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 608  
 ; LENGTH: 858  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-770-445-608

Alignment Scores:  
 Pred. No.: 1,4e-51 Length: 858  
 Score: 663.00 Matches: 119  
 Percent Similarity: 70.87% Conservative: 44  
 Best Local Similarity: 51.74% Mismatches: 67  
 Query Match: 23.21% Indels: 0  
 DB: 9 Gaps: 0

US-10-019-783-2 (1-551) x US-09-770-445-608 (1-858)

Qy 320 CysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeu 339  
 Db 858 TGTGGGAATGTTTCTCTCGCCACATCTTCAAAAGATTGCAGACGCTTGCAGCTT 799  
 Qy 340 GlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPhe 359  
 Db 798 GGTATACCTGTGATCCGACGCAAGCTCTATGACCATTTTGCCTTTTGGGATAAACCATTT 739  
 Qy 360 IleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLys 379

Db 738 GTGTCATGGCAGATTGGCAGACTAGTCCCTGTTATTGCTTAGGTGCTATATCTAA 679  
Qy 380 SerTrpIleValProGlyTTPArgLeuGlyTTPValAlaValTyrAspProArgLysIle 399  
Db 678 AGATGGTTTGTCTCGATGGAGACTTGGTTGGATGGTGAATCTCTTGACCTCATGGCATC 619  
Qy 400 LeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPro 419  
Db 618 ATGAAGAAGATTGGGTTGTTTCAGACTTCTATCAATGTTGTCACATGTGCAGCATCCT 559  
Qy 420 AlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePhe 439  
Db 558 GCAAGTTTATTCAGGAGCAATGCCCTGATATCATTTGGGAATACAAAGGAAGAGTTCTTC 499  
Qy 440 LysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGlu 459  
Db 498 TCATCAAAACTTGAATGGTGAATAATGTGCAGAGATTGTTATGAGGAGCTTATGAAG 439  
Qy 460 AsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeu 479  
Db 438 ATTCTTGCATCACTTGGCCCTGCBAACCCGAGGGCTCAATGTTCAAGATGGTGAAGTTA 379  
Qy 480 AsnLeuHisLeuLeuGluGluIleAspAspIleAspPheCysCysLysLysLeuAlaLys 499  
Db 378 AACTTTTCACACTCTCGAAGATATCAGTGATGATTTGGACTTCTCTCCAAAGCTGGCTAAA 319  
Qy 500 GluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTTPValArgile 519  
Db 318 GAGGAATCTATGATCATCTACCAGTCAAGCTGTGGCTTAAGAACTGGCTACGTATC 259  
Qy 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys 539  
Db 258 ACCTTTCAGTTGAGCTTGAGCTTCTCATAGAAGGGTTTCCAGGCTAAAGAACTTTACT 199  
Qy 540 GlnArgAsnLysLysArgAsnSerSerAsp 549  
Db 198 GAGAGACACTCCAAATCAATCAGCCATGAGAC 169

RESULT 4  
US-09-919-039-379/c  
; Sequence 379, Application US/09919039  
; Publication No. US20030108871A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919, 039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/222,113  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 379  
; LENGTH: 2712  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20030108871A1 335648.1c  
; NAME/KEY: unsure  
; FEATURE:  
; LOCATION: 236-590  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-919-039-379

Alignment Scores:  
Pred. No.: 4,49e-58 Length: 2712  
Score: 637.50 Matches: 133  
Percent Similarity: 54.77% Conservative: 85  
Best Local Similarity: 33.42% Mismatches: 155  
Query Match: 22.32% Indels: 25  
DB: 11 Gaps: 6

US-10-019-783-2 (1-551) x US-09-919-039-379 (1-2712)  
Qy 167 ProArgPro-----ValLeuProLeuAlaHisGlyAspProSerValPheProAla 183  
Db 2418 CCAATATCCAAACAAACCATGATTTCCTCTGCTCCATTTGGGGACCTCTGTTTGGAAAC 2359  
Qy 184 PheArgThrAlaValGluAlaAspAlaValAlaAlaValArgThrGlyGlnPhe 203  
Db 2358 CTGCTCAGACACCTGAAATGTTACCCAGGCAATGAAAGATGCCCTCGGACCTCGGCAATAT 2299  
Qy 204 AsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu----- 221  
Db 2298 AATGCTATGCCCATCCATCGCTTCTATCCATCGGGAGGAGATTGCTTCTTATTATC 2239  
Qy 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241  
Db 2238 CACTGCTCTGAGGCA-----CCCCTAGAGCTAAGGACGTCATCTCACAAGTGGC 2188  
Qy 242 GlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261  
Db 2187 TGCAGCCAAAGCTATTGACCTTTGTTAGCTGTGTTGGCC---AACCCAGGGCAAAACATC 2131  
Qy 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281  
Db 2130 CTGGTTCCAGACCTGGTTCTCTCTACAAGACTCTGGCTGAGTCTATGGAAATTCAG 2071  
Qy 282 ValArgHisPheAspLeuIleProAspLysGlyTTPGluIleAspIleAspSerLeuGlu 301  
Db 2070 GTCAAACTCTACAATTTGTCGAGAGAAATCTGGGAAATTCAGCTGAAACAACTGGAA 2011  
Qy 302 SerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGly 321  
Db 2010 TATCTAATTTGATGAAGACACCTTGTCTCATGTGCAATATCCATCAACCCCTGTGG 1951  
Qy 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341  
Db 1950 TCAGTGTTCAGCAAAAGCTCATCTTCAGAAAGATTCTGGCAGTGGCTGCACGGCAGTGTGC 1891  
Qy 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361  
Db 1890 CCCATCTTAGCTGATGATCTATGAGACATGGTGTGTTTCGGATGTGCAATATGAACA 1831  
Qy 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLysSerLysTrp 381  
Db 1830 CTGGCCACCTCAGCACCATGTCCTCCATCTCTCTGTGGAGGCTGCCCCAGGCTGG 1771  
Qy 382 IleValProGlyTTPArgLeuGlyTTPValAlaValTyrAspProArgLysIleLeuGln 401  
Db 1770 CTGGTTCTCTGGCTGGAGTGGCTGGATCTCATGACCGAGAGACATTTTGGC 1711  
Qy 402 Glu-----ThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 415  
Db 1710 AATGATCCGAGATGGGCTGGTGAAGCTGAGTCAGCGCATTTTGGGA----- 1663  
Qy 416 SerThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLys 435  
Db 1662 -----CCCTGTACCATTTGCCAGGGAGCTCTGAAAAGCATCTATGTCGCCACCCG 1612  
Qy 436 GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTrpLys 455  
Db 1611 GGAGAGTTTTACCAACACTCTGACCTTCTCAAGTCCATGCTGATCTCTGTTATGG 1552  
Qy 456 GlnIleLysGluAsnLysTyrIleThrCysProHisLysProGlySerMetPheVal 475  
Db 1551 GCCTTGGCTGCCATCCCTGAGACTCCGG---CCAGTCCGCCCTCTGGGGCTATGACCTC 1495  
Qy 476 MetValLysLeuAsnLeuHisLeuLeuGluLeuAspAspIleAspPheCysCys 495  
Db 1494 ATGGTTGGAATTCAGATGGAACAAATTTCCAGAAATTTGAGACAGATGTGGAGTTCCAGGAG 1435  
Qy 496 LysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsn 515  
Db 1434 CGGTTAGTTGCTGAGCAGTCTGTCCACTGCTCCAGCAACCGTCTTGGTGGTACCGGAAT 1375



QY 516 TtpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle 535  
Db 1374 TTCTCCGAGTGTCTATCATCAGTCCCGAGGTGATGATCTCGAGCGGTGCGACCGGATC 1315  
QY 536 LysSerPheCysGlnArgAsnLys-----LysArgAsnSerSerAspAspCys 551  
Db 1314 CAGGAGTTCTGTGAGCAGCACTACCATTTGCTGTAAGGCGAGCGAGGAGTGT 1261

## RESULT 5

US-09-880-107-3720  
; Sequence 3720, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3720  
; LENGTH: 2754  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X52520  
US-09-880-107-3720

Alignment Scores:  
Pred. No.: 4,598-58 Length: 2754  
Score: 637.50 Matches: 133  
Percent Similarity: 54.77% Conservative: 85  
Best Local Similarity: 33.42% Mismatches: 155  
Query Match: 22.32% Indels: 25  
DB: 10 Gaps: 8

US-10-019-783-2 (1-551) x US-09-880-107-3720 (1-2754)

QY 167 ProArgPro-----ValLeuProLeuAlaHisGlyAspProSerValPheProAla 183  
Db 295 CCAATCCAAACAAACCAATGATTTCCCTGCTCCATTTGGGACCTACTGTGTTTGGAAAC 354  
QY 184 PheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArgThrGlyGlnPhe 203  
Db 355 CTGCCTACAGACCTCGAAGTTACCCAGCAATGAAGATGCCCTGCGACTCGGGCAAAATAT 414  
QY 204 AsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu----- 221  
Db 415 AATGCTATGCCCATCCATCGCTTCTATCCAGTCCGGAGGATGCTCTTATATAC 474  
QY 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241  
Db 475 CACTGCTCTCTGGGCA-----CCCCTAGAAGCTAAGGACGTCATTCTGACAAGTGGC 525  
QY 242 GlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261  
Db 536 TGCAGCCAGCTATGACCTTGTAGTCTGTTGGCC-----AACCAGGGCAGAACATC 582  
QY 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281  
Db 583 CTGTTCCAGACCTGGTTCTCTCTACAGACTCTGCTGAGTCTATGGGAATTGAG 642  
QY 282 ValArgHisPheAspLeuProAspLysGlyTrpGluLeuAspIleAspSerLeuGlu 301  
Db 643 GTCAAACTCTACAATTTGTTCCAGAGAAATCTTTGGGAAATTTGACCTGAAACCACTGGAA 702

QY 302 SerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGly 321  
Db 703 TATCTAATTGATGAAGACAGCTGTCTCATTGTCATTAATCCATCAACCCCTGTGGG 762  
QY 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyLe 341  
Db 763 TCAGTGTTCAGAAACCTCATCTTCAGAGATCTTGGCAGTGGCTGCACGGCAGTGTGTC 822  
QY 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361  
Db 823 CCCATCTTAGCTGATGAGATCTATGGAGACATGGTGTTCGGATTCGAATATGAACCA 882  
QY 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTyr 381  
Db 883 CTGGCCACCTTCAGCACCGGATGCCCATCTCTCTGTGGAGGCTGGCAAGCGGTGG 942  
QY 382 IleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGln 401  
Db 943 CTGGTTCCTGGCTGGAGTTGGCTGGATCCTCATTCATGACCGAAGAGACATTTTGGC 1002  
QY 402 Glu-----ThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 415  
Db 1003 AATGAGATCCGAGATGGGCTGGTGAAGCTGAGTCAGCGCATTTTGGCA----- 1050  
QY 416 SerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLys 435  
Db 1051 -----CCCTGTACCATTTCCAGGGAGCTCTGAAAGCATCTTATGTCACCCCG 1101  
QY 436 GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLys 455  
Db 1102 GGAGAGTTTACCACCAACTCTGAGCTTCTCAAGTCCCAATCTGATCTCTGTATGG 1161  
QY 456 GluIleLysGluAsnLysTyrIleThrCysProHisLysProGlnGlySerMetPheVal 475  
Db 1162 GCGTTGGCTGCCATCCCTGGACTCCGG-----CCAGTCCGCCCTTCTGGGCTATGTACCTC 1218  
QY 476 MetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspIleAspPheCysCys 495  
Db 1219 ATGTTGGAAATTGAGATGGAAACATTTCCAGAAATTTGAGAACATGTGGAGTTACGGAG 1278  
QY 496 LysLeuAlaLysGlnGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsn 515  
Db 1279 CGGTAGTTGCTGAGCAGCTGTGTCACCTGCTCCAGCAACGCTGTTGAGTACCCGAAT 1338  
QY 516 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle 535  
Db 1339 TTCTCCGAGTGTCTATCATCAGTCCCGAGGTGATGATGCTGGAGCGGTGCACCCGATC 1398  
QY 536 LysSerPheCysGlnArgAsnLys-----LysArgAsnSerSerAspAspCys 551  
Db 1399 CAGGAGTTCTGTGAGCAGCACTACCATTTGCTGTAAGGCGAGCGAGGAGTGT 1452

## RESULT 6

US-10-198-846-11998  
; Sequence 11998, Application US/10198846  
; Publication No. US2003099974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11998

4

```

QY 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241
Db 487 CACTGCTCTGGAGCA-----CCCTAGAGCTAAGAGCTCATTTGCAAGTGGC 537
QY 242 GlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261
Db 538 TGCAGCAAGACTATTGACCTTTGTTAGCTGTGTGGCC---AACCCAGGGCAAAACATC 594
QY 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281
Db 595 CTGGTTCCAGAGCTGGTTCTCTCTACAGACTCTGGCTGAGCTATGCGGAATTGAG 654
QY 282 ValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGlu 301
Db 655 GTCAAACTCTACAATTGTTGCCAGAGAAATCTGGGAATTCAGCTGAAACAATGGA 714
QY 302 SerIleAlaAspLysAsnThrThrAlaMetValIleAsnProAsnAsnProCysGly 321
Db 715 TATCTAATTGTAAGAAAGACAGCTTGCTCATTTGTCATTAATCCATCAACCCCTGTGG 774
QY 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341
Db 775 TCAGTGTTCAGCAACGCTCATCTTCAGAGATCTGGCAGTGGCTGCAGCGAGTGTCTC 834
QY 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361
Db 835 CCATCTCTAGCTGATGAGATCATGAGACATGGTGTTCGGATTGCAATATGCAACCA 894
QY 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTrp 381
Db 895 CTGGCCACCTCAGCACCGATGCTCCCATCCCTGCTGTGGAGGCTGGCCAAAGCGTGG 954
QY 382 IleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGln 401
Db 955 CTGGTCTCTGGCTGAGGTGGCTGGATCCTCATTCATGACCGAGAGACATTTTGGC 1014
QY 402 Glu-----ThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 415
Db 1015 AATGATCCCGAGATGGGCTGGTGAAGCTGAGTCAGCGCATTTTGGGA----- 1062
QY 416 SerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLys 435
Db 1063 -----CCTGTACCATTTGTCAGGAGCT----- 1086
QY 436 GluAspPhePheLysAlaIleGlyLeuLeuLysGluSerSerGluIleCysTyrLys 455
Db 1087 -----CTGCTCAAGTCCAAATGCTGATCTCTGTTATGG 1119
QY 456 GlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 475
Db 1120 CGGTGGTGGTCCATCCCTGGACTCGG---CCAGTCCCGCTCTGGGGCTATGATCCTC 1176
QY 476 MetValLysLeuAsnLeuHisLeuLeuGluIleAspAspAspIleAspPheCysCys 495
Db 1177 ATGGTTGAATAGATGAACATTTCCAGAAATTTGAGAACGATGTGGAGTTCACGGAG 1236
QY 496 LysLeuAlaLysGluLysValIleLeuLysProGlySerValLeuGlyMetAlaAsn 515
Db 1237 CGGTTAGTGTGAGCAGTCTGTCCATCGCTCCAGCAACGCTGTTGAGTACCCGAA 1296
QY 516 TrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle 535
Db 1297 TTCATCCAGTGGTTCATCAGAGTCCCGAGGTGATGCTGGAGGGCTGCAGCGCGATC 1356
QY 536 LysSerPheCysGlnArgAsn 542
Db 1357 CAGGAGTCTCTGTGAGCAGCAC 1377

```

## RESULT 8

US-10-198-846-8896/c  
 ; Sequence 8896, Application US/10198846  
 ; Publication No. US2003009974A1  
 ; GENERAL INFORMATION:

APPLICANT: Lillie, James  
 APPLICANT: Xu, Yongyao  
 APPLICANT: Wang, Youzhen  
 APPLICANT: Steinmann, Kathleen  
 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
 FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 THERAPY OF BREAST CANCER  
 FILE REFERENCE: MEI-049  
 CURRENT APPLICATION NUMBER: US/10/198,846  
 CURRENT FILING DATE: 2002-07-18  
 PRIOR APPLICATION NUMBER: 60/306,220  
 PRIOR FILING DATE: 2001-07-18  
 NUMBER OF SEQ ID NOS: 14084  
 SOFTWARE: fastseq for Windows Version 4.0  
 SEQ ID NO 8896  
 LENGTH: 870  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 24, 687, 693, 705, 723, 757, 768, 814, 826  
 OTHER INFORMATION: n = A,T,C or G  
 US-10-198-846-8896

## Alignment Scores:

Pred. No.: 3 87e-33 Length: 870  
 Score: 399.00 Matches: 101  
 Percent Similarity: 52.10% Conservative: 48  
 Best Local Similarity: 35.31% Mismatches: 101  
 Query Match: 13.97% Indels: 37  
 DB: 14 Gaps: 7

US-10-019-783-2 (1-551) x US-10-198-846-8896 (1-870)

```

QY 164 GluLysGlyProArgPro-----ValLeuProLeuAlaHisGly----- 176
Db 849 GAACCAATCCCAACCAAAANAATGTTTCCNTGGCCCAATGGGGGAGCCCTACC 790
QY 177 -----AspProSerValPheProAlaPheArgThra 187
Db 789 TGGTGTGGGAAACCTGGCTACAGACCCNTTAAAGTTTCC----- 745
QY 188 ValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTyrPro 207
Db 744 ---CAGGCCAAATAAAAGATGCCNTGGAACTCGGGCAAAANAAAGGGCTANGGCC 688
QY 208 AlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyVal 227
Db 687 NATCCATCGGCTTCCCTA-TCCAACTCGGGAAGAGATTGTTCTTTATTACCACTGT 629
QY 228 Pro-----TyrMetLeuSerAlaAspValPheLeuThr-AlaGlyGlyThrGlnAl 245
Db 628 CCTTGAAGGGCACCCCTAGAGCTAAGGAGCTATTCTGCAAGTGGTGCAGCCAGC 569
QY 245 alleGluValIle-IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProA 265
Db 568 TATTCAGCTTTGTTTGTAGCTGTGTGGCC---AACCCAGGGCAAAACATCTGGTTCCAA 512
QY 265 rgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisP 285
Db 511 GACCTGTTTCTCTCTACAGACTCTGGCTGAGTCTATGGAAATTGAGGTCAACTCT 452
QY 285 heAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaA 305
Db 451 ACAATTGTTGGCAGAGAAATCTTGGAAATTTGACCTGAAACAACATGGAATATCTAATG 392
QY 305 splysAsnThrThrAlaMetValIleAsnProAsnAsnProCysGlySerValTyrS 325
Db 391 ATGAAAGACAGCTGTCTCATTTGCAATAATCCATCAACCCCTGGGTGAGTGTCA 332
QY 325 ertYrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleA 345
Db 331 GCAACGTCATCTTCAGAGAGATCTTCAGAGTGGCTCCAGCGGAGTGTGTCCTCCACTTAG 272

```

QY 345 laAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValp 365  
 Db 271 CTGATGAGATCTATGGAGACATGGTGTTCCTCGGATTGCAATATGAACCACTGGCGGCC 212  
 QY 365 heGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTyrIleValProG 385  
 Db 211 TCAGCACCAGTGTCCCAATCTGCTGTGTGAGGGCTCGCAACGGCTGGCTGTCTG 152  
 QY 385 lyTPrArgLeuGlyTPrValAlaValTyrAspProArgLysIleLeuGlnGlu----- 402  
 Db 151 GCTGAGGCTGGCTGGATCTCTCATGACCAAGAGACATTTTGGCAATGAGATCC 92  
 QY 403 -----ThrIysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAsp 419  
 Db 91 GAGATGGCTGGTGAAGCTGAGTCAGCGCAIT-----TTGGGACCTGTACCTCGG 41  
 QY 419 roAlaThrPheIle 423  
 Db 40 CGGCCACCGCGGTG 27

## RESULT 9

US-10-156-761-4891

; Sequence 4891, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; PRIOR FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 4891

; LENGTH: 1224

; TYPE: DNA

; ORGANISM: Streptomyces avermitilis

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1224)

US-10-156-761-4891

Alignment Scores:

Pred. No.:	1,17e-28	Length:	1224
Score:	359.50	Matches:	110
Percent Similarity:	46.71%	Conservative:	89
Best Local Similarity:	25.82%	Mismatches:	175
Query Match:	12.53%	Indels:	52
DB:	14	Gaps:	18

US-10-019-783-2 (1-551) x US-10-156-761-4891 (1-1224)

QY 141 ValLeuAlaAlaThrGlyValAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160  
 Db 34 GTCTCGCCCGAGTGGCGCGATCTCCGAGTCGCGCACCTCGCGGTGATGCCAAGGCC 93  
 QY 161 SerValGlnGluGlyGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180  
 Db 94 AAGGCCCTCAAGCGCCCGCGAGCTCGGTGATCGGCTTCGCGCGCGGTGAGCCGAC--- 150  
 QY 181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThr 200  
 Db 151 TTCCCGACCCCGGACTACATCTCCAGGCC-----GCCATCGAGGCTGTCTGNAACCCG 204

QY 201 GlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 220  
 Db 205 AAGTACCACCGCTACACCGCGCGGC---GGCCTGCCGAGCTGAAGCGCGCATCGCC 261  
 QY 221 GluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAla 240  
 Db 262 GCGAAGACGCTCCGCGACTCGGGCTACGAGGTGGAGCTCGCAGGTCTCTGTCACCAAC 321  
 QY 241 GlyGlyThrGlnAlaIle---GluValIleIleProValLeuAlaGlnThrAlaGlyAla 259  
 Db 322 GCGCGACAGCAGCGCATCTACAGCGCTTCGCCGCAATCTC-----GACCCGGCGAC 375  
 QY 260 AsnIleLeuLeuProArgProGlyTyrProAsnTyr---GluAlaArgAlaAlaPheAsn 278  
 Db 376 GAGGTATCTGCTCCCGCGCGTACTGACCACTACCCGGAGTCGATCCGCTCGCCGCGC 435  
 QY 279 ArgLeuGluValArgHisPheAsnLeuIleProAspLys-----GlyTPrGluIleAsp 296  
 Db 436 GGTGTCCCGCTC-----GACGTCTGTCGCGCAGAGACACCGGCTACCGGTCTCG 486  
 QY 297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 316  
 Db 487 GTCGAGCAGCTGAGGCGCGCCGCGACCGAGAACCAAGGTCTGCTCTCTCTCCCGC 546  
 QY 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336  
 Db 547 TCCAAACCCGACGCGCGGTCTACACCGCGAGATCGAGGATCGGCGCTGGGCC 606  
 QY 337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356  
 Db 607 GCGGAGAAGGCTGTGGGTCTCTGACCGAGAGATCTACGAGCCTCTCTACGGGAC 666  
 QY 357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSer----- 373  
 Db 667 GCGGAGTTCCTCCTCCTCGCGTG-----GTCGTGCGCGAGCTGCGGACCAAGTGC 717  
 QY 374 -----IleGlySerLeuSerLysSerThrIleValProGlyTPrArgLeuGlyTPrVal 391  
 Db 718 ATCGTGTCAACGGTGTGCGGAAGACGTACGCGGATCGCGGTGGCGGTGGCTGG--- 774  
 QY 392 AlaValTyrAspProArgLysIleLeuGln-----GluThrLysIleSer 406  
 Db 775 ---GTCATCGGTCCCAAGACGCTCATCAAGCGCGCGAGAACCTCCAGCTCGACGCCACG 831  
 QY 407 ThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAla 426  
 Db 832 TCGAACGTCTCAACGTGCGCGAGTCTGCGCGCTCGCGCGCC-----GTGTCGCGCAT 885  
 QY 427 LeuProGlnIleLeuGluAsnThrLysGluAspPhe-----PheLysAlaIleIle 443  
 Db 886 CTGACGCGCTCGCGAG---ATGCGGAGGCTTCGACCGCGCGCGCAAGACCATCTGTG 942  
 QY 444 GlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIle 463  
 Db 943 CGGATCTCAACGAGATCGCGCGC-----GTC-----GTC 969  
 QY 464 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 483  
 Db 970 CTGTGCGCG---GAGCCCGAGGCGGCTTCTACGCGTACCCCTCGGTGAAGCGGTGCTC 1026  
 QY 484 LeuGluGluIle-----AspAspAspIleAspPheCysCysLysLeuAla 498  
 Db 1027 GGCAGGAGATCCGCGCGCAAGCGCGCGAGACACGCTGAGCTGGCGCGCTGATCTCTG 1086  
 QY 499 LysGluLeuSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArg 518  
 Db 1087 GAGAGCGCGAGTCCGCTGTCTCCCGCGAGGCTTCGCGCACCGCGGCTATCTGGCG 1146  
 QY 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle-LysSerPh 538  
 Db 1147 CTGTCTACGCCCTGGTGACGAGGATCTGTGTCGAGGCGGTGAGCGGGATTGAGAGCTG 1206  
 QY 538 eCysGlnArgAsnLys 543

```

Db 1207 CTGTCCGAGCGGAAG 1222
RESULT 10
US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIEN, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Alignment Scores:
Pred. No.: 8,77e-23 Length: 9025608
Score: 359.50 Matches: 110
Percent Similarity: 46.71% Conservative: 89
Best Local Similarity: 25.82% Mismatches: 175
Query Match: 12.59% Indels: 52
DB: 14 Gaps: 18

US-10-019-783-2 (1-551) x US-10-156-761-1 (1-9025608)
QY 141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160
Db 5964250 GTCTCCGCCGAGTCGGCGCGATCTCCGAGTCGCCACCCCTCGCGCGTGATGCCAAGGCC 5964191
QY 161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
Db 5964190 AAGGCCCTCAAGCGCGCGGACGTCGGGTGATCGGCTTCGCGCGCGGTGACCGGAC--- 5964134
QY 181 PheProAlaPheArgThrAlaValGlnAlaGluAspAlaValAlaAlaAlaValArgThr 200
Db 5964133 TTCCCGACCCGGACTACATCGTCAGGCC-----GCCATCGAGGCGCTGCTCGAACC 5964080
QY 201 GlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 220
Db 5964079 AAGTACGACCCGTACACCGCGCGCGC---GGCTGCGCGAGCTGAAGCGCGCATCGCC 5964023
QY 221 GluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAla 240
Db 5964022 CGCAAGACCGCTCCGCGATCGGGGTACGAGTGGACGCGCTCGCAGGTCCTGGTCAACAC 5963963
QY 241 GlyGlyThrGlnAlaIle---GluValIleIleProValLeuAlaGlnThrAlaGlyAla 259
Db 5963962 GCGCGCAGCAGGCGATCTACAGCGCTTCGCCCGGATCTCTC-----GACCCGGGGGAC 5963909
QY 260 AsnIleLeuLeuProArgProGlyTyrProAsnTyr---GluAlaArgAlaAlaPheAsn 278
Db 5963908 GAGGTATCGTCCCGCGCGCGCTACTGGACCACTACCGGAGTCGATCCGCTCTCGCGCGC 5963849
QY 279 ArgLeuGluValArgHisPheAspLeuIleProAsnLys-----GlyTyrGluIleAsp 296

```

## RESULT 11

```

US-09-905-173-39
; Sequence 39, Application US/09905173
; Patent No. US20020132295A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay M.
; APPLICANT: WARREN, Patrick V.
; APPLICANT: SWANSON, Ronald V.
; APPLICANT: NATHUR, Eric J.
; TITLE OF INVENTION: ENZYMES HAVING TRANSAMINASE AND AMINOTRANSFERASE ACTIVITY AND MEET
; FILE REFERENCE: DIVER1240-7
; CURRENT APPLICATION NUMBER: US/09/905,173

```

```

Db 5963848 GGTGTCCCGCTC-----GACGTCTGCCGACGAGACCACCGCGCTACCGGTCTCG 5963798
QY 297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 316
Db 5963797 GTCGAGCAGGTGGAGCGCGCCGACGAGAACACCAAGGTGTCTCTTCTCTCCCG 5963738
QY 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336
Db 5963737 TCCAAACCCGAGCGCGCGGTCTACACCGGAGAGATCAGGAGATCGGCCGCTGGGCC 5963678
QY 337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
Db 5963677 GCCGAGAAGGCGCTGTGGTCTTACCGACGAGATCTACGAGCACCTCGTCTACGGGAC 5963618
QY 357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSer----- 373
Db 5963617 GCCGAGTTCACCTCCCTGCCGGTG-----GTCGTGCCGAGTGGCGCAAGATGC 5963567
QY 374 -----IleGlySerLeuSerLysSerTyrIleValProGlyTyrArgLeuGlyTyrVal 391
Db 5963566 ATCGTGTCTCAACGTCGCGGAGAGCGTAGCGATGACCGGCTGGCGGTGGGCTGG--- 5963510
QY 392 AlaValTyrAspProArgLysIleLeuGln-----GluThrLysIleSer 406
Db 5963509 ---GTCATCGTCTCCCAAGGACGTCATCAAGCGCGCGACCACTCCAGTCCGACGCACG 5963453
QY 407 ThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAla 426
Db 5963452 TCGAAGCTCTCCAACTGGCGGAGTGGCGGCTCGCGCTCGCGCC-----GTGTCCGGCGAT 5963399
QY 427 LeuProGlnIleLeuGlnAsnThrLysGluAspPhe-----PheLysAlaIleIle 443
Db 5963398 CTGACGGCGCTCGCGGAG---ATGCGGAGGCGCTTCGACGCGCGCGCAAGACCATCGTG 5963342
QY 444 GlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIle 463
Db 5963341 CGGATGCTGAACGAGATCGCGGC-----GTC 5963315
QY 464 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 483
Db 5963314 CTGTGCGCG---CAGCCCGAGGCGGCTTCTAGCGGTACCCCTCGGTGAAGGCGCTGC 5963258
QY 484 LeuGluGluIle-----AspAspAspIleAspPheCysCysLysLeuAla 498
Db 5963257 GGCAAGAGATCGCGGCAAGCGCGCGAGGACACGCTCGAGCTCGCGCTGATCTCTG 5963198
QY 499 LysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyrValArg 518
Db 5963197 GAGGAGCGCGGTGCGGTGCTCCCGCGAGGCGCTTCGGACGCGCGGTATCTCGCG 5963138
QY 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIle-LysSerPh 538
Db 5963137 CTGTCTGCGCTCGGTGACGAGGATCTGTGAGGCGGTGAGCGGATTCAGAGCTG 5963078
QY 538 eCysGlnArgAsnLys 543
Db 5963077 CTGTCCGAGCGGAAG 5963062

```

; CURRENT FILING DATE: 2001-07-12  
 ; PRIOR APPLICATION NUMBER: US 09/412,184  
 ; PRIOR FILING DATE: 1999-10-04  
 ; PRIOR APPLICATION NUMBER: US 09/389,537  
 ; PRIOR FILING DATE: 1999-09-02  
 ; PRIOR APPLICATION NUMBER: US 08/646,590  
 ; PRIOR FILING DATE: 1996-05-08  
 ; PRIOR APPLICATION NUMBER: US 08/599,171  
 ; PRIOR FILING DATE: 1996-02-09  
 ; PRIOR APPLICATION NUMBER: US 09/481,733  
 ; PRIOR FILING DATE: 2000-01-11  
 ; PRIOR APPLICATION NUMBER: US 09/069,226  
 ; PRIOR FILING DATE: 1998-04-27  
 ; NUMBER OF SEQ ID NOS: 40  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 39  
 ; LENGTH: 1185  
 ; TYPE: DNA  
 ; ORGANISM: Aquifex  
 ; US-09-905-173-39

Alignment Scores:  
 Pred. No.: 1,936-27 Length: 1185  
 Score: 348.00 Matches: 105  
 Percent Similarity: 47.72% Conservative: 83  
 Best Local Similarity: 26.65% Mismatches: 166  
 Query Match: 12.18% Indels: 40  
 DB: 10 Gaps: 15

US-10-019-783-2 (1-551) x US-09-905-173-39 (1-1185)

QY	158	IleSerAlaSerValGlnGlu-----LysGlyProArgProValLeuProLeuAla	174
DB	58	ATAACCGAAGAAAGCAAGATTAAAGGCTAAGAGTG-----GAGCTATAGGTTTTGGA	114
QY	175	HisGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaVal	194
DB	115	GGGGGAGAACCTGAC-----TTCCGACACACCCGACTTCATAAAGAAAGACCTGT	162
QY	195	AlaAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAla	214
DB	163	ATAAGGGCTTTTAAGGGAAGGAAGAAACCAAG---TACGCTCCCTCCGGGGAATACCAAG	219
QY	215	AlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAsp	234
DB	220	CTCAGAGAAGCTATAGCTGAAAACTACTGAAGAAACAAAGTTGAGTACAAACCTTCA	279
QY	235	AspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAla	254
DB	280	GAGATAGTCGTTTCCCGCAGGAGCGAAATGGTTCTTCTCCATATATTCATGCTATACTG	339
QY	255	GlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArg	274
DB	340	GAC---GAAGGAGACAGAGTTTACTAGCTTACTGCCCTTACTGGGTAACTTACCCGAA--	393
QY	275	AlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIlePro-----	289
DB	394	-----CAGATAAGTTCTTCGGAGGGGTTCCGTTGAGGTTCTCTCTA	435
QY	290	-----AspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsn	307
DB	436	AAGAAAGAGAAAGGATTTCATTAAGTCTGGAAGATGCAAGAAAGGTTACCGAGAGA	495
QY	308	ThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAsp	327
DB	496	ACAAAAGCTATAGTCATAAATCTCCGAACACCCCACTGCTGCTGTTTACGAAGAGG	555
QY	328	HisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGlu	347
DB	556	GAACTTAAGAAATAGCGGGGTTTTCGCTGGAGAGGGGCAITTCATTAATTTCCGATGAG	615
QY	348	ValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly---	366

DB	616	TCCTATGAGTACTTCGTTTACGGTATGCAAAATTTGTTAGCCCTGCTCTTTCTCGGAT	675
QY	367	-----HisIleThrProValLeuSerIleGlySerLeuSerLysSerTrpIleVal	383
DB	676	GAAAGTAAAGAACATAAC-----TTACGGTAAACGCCCTTTTCAAGAGCTATTCCATG	729
QY	384	ProGlyTrpAspGlyLeuGlyTrpValAlaValTyrAspProArgLysIleLeuGlnGluThr	403
DB	730	ACTGGTTGGCAATAGATTATGACGTCGCCGAA-----GAGTACGCA	774
QY	404	LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIle	423
DB	775	AAAGTATAGCGAGTCTT---AACAGCCAGAGTGTTC---AACGTCACTACCTTTGCC	828
QY	424	GlnAlaAlaLeuProGlnIleLeuGluAsnThrLys---GluAspPhePheLysAlaIle	442
DB	829	CAGTATGGAGCTCTTGGAGCCCTTGAATAATCAAAGTCTAAAGATTTCGTAACGAAATG	888
QY	443	IleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyr	462
DB	889	AGAAATGCTTTTGAAGGAGGAGGATACGGCTGTAGAAGAGCTTCTTAAATTCACGT	948
QY	463	IleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHis	482
DB	949	ATGGATGTGTA---AAACCCGAGGTGCTTTTACATATTTCCGAGACTTCTCCGCTTAC	1005
QY	483	LeuLeuGluGluIleAspAspIleAspPheCysLysLysLeuAlaLysGluGluSer	502
DB	1006	---GCTGAGAAACTGGTGGTGTGATGTGAACCTCTCGGAGTTCCTTCTGAAAAGGCTAAG	1062
QY	503	ValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIleThrPheAla	522
DB	1063	GTTGCGTGTTCCTCCGCTTCGGAGCTCCCGGATTTTTCGAGCTTCTTACGCC	1122
QY	523	CysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys	536
DB	1123	CTTTCGAGAAAGACTCGTTGAGGATTAAGGAGATAAAG	1164

RESULT 12  
 US-09-738-626-1  
 ; Sequence 1, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; PRIOR FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-06-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 3309400  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 ; US-09-738-626-1

Alignment Scores:  
 Pred. No.: 1,666-21 Length: 3309400

[illegible]

Db 61 CCTGGATGGAGCGTTGGTTGGTTGGTGCACAAATGATCCATCTGGCAGCTTTTGAATCCA 120  
 Qy 404 LysileSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIle 423  
 Db 121 AAGGTAGATGAGCGCTTTAAAGAGTACTTTGATCGTTGGAGGTCCAGCCAGCTTCATC 180  
 Qy 424 GlnAlaLalaProGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleIle 443  
 Db 181 CAGCAGCTGCTACCTCAGATTAATGAGCATACTGAAAGAGTTTCTTCAAGAAACCAT 240  
 Qy 444 GlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIle 463  
 Db 241 GATAATTGAGCGATGTTGCACATATATGTTGTAAGAGTTGAAGGATATTCATACATT 300  
 Qy 464 ThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeu 483  
 Db 301 ATATGCCCTTATAACACCAAGAGTCCATGGCTATGATGGTGAACATAAACCTTTTCATT 360  
 Qy 484 LeuGluGlu 486  
 Db 361 CTGGAGGAT 369

## RESULT 14

US-09-738-626-3134  
 ; Sequence 3134, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIKO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: Patentin ver. 3.0  
 ; SEQ ID NO 3134  
 ; LENGTH: 1257  
 ; TYPE: DNA  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3134

## Alignment Scores:

Pred. No.: 1,968-26 Length: 1257  
 Score: 339.00 Matches: 108  
 Percent Similarity: 45.52% Conservative: 85  
 Best Local Similarity: 25.47% Mismatches: 173  
 Query Match: 11.87% Indels: 58  
 DB: 10 Gaps: 16

US-10-019-783-2 (1-551) x US-09-738-626-3134 (1-1257)

Qy 151 IleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArg----- 168  
 Db 70 ATGAAGGACGTGTGTGTACGAGATCGGTGCCCGGTGGCGGGGAGCGGCGCATGGAG 129  
 Qy 169 -----ProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 184  
 Db 130 CTTGTATGGGCATACATCTTAAAGCTCAACCGGGAATTCAGCGGTGTTCC----- 180

Qy 185 ArgThrAlaValGluAlaGluAspAlaVal-----AlaAlaAlaValArgThr 200  
 Db 181 -----GATTCGATGCCCCGACGCTGATTATCGTGACATGATCGCCCAACCTTCCAACT 234  
 Qy 201 GlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 220  
 Db 235 TCCCAAGGGTATTCCACCTCCAAAGGCATT---ATTCCGGCCCGCGAGCAGCTGGTCAAC 291  
 Qy 221 GluHis---LeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThr 239  
 Db 292 CCTCAGAGTTGTCGCCGATTTCCCCAC---TTCGATGTTGATGATGTTCTTAGGC 348  
 Qy 240 AlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAla 259  
 Db 349 AACGGTGTCTCAGAACTAATCACCATGACCAACCAAGCCTCTCAACGAC---GGCGAT 405  
 Qy 260 AsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArg 279  
 Db 406 GAAGTTCTTATCCCGACCGGACTACCCACTGTGGACTGCGCAACTCCTCGCTGGT 465  
 Qy 280 LeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspLeuSer 299  
 Db 466 GGTAAAGCCTGTGCTACTCTCTGTGATGAGGAAGATGACTGGAACCCATCCATCAAGAC 525  
 Qy 300 LeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnPro 319  
 Db 526 ATCAAGTCCAAAATCTCAGAGAAAACCAAGCTATTGTGGTGATCAACCCCAACACCCC 585  
 Qy 320 CysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeu 339  
 Db 586 AGCGAGCTGTCTACCCGCGCGGTGTGGACAATCGTCGATGTCAGCGAGCAT 645  
 Qy 340 GlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPhe 359  
 Db 646 GACCTGCTGATTTGGCCGATGAAATACGACCGCATTTCTACGATGATGATGATGATGAT 705  
 Qy 360 IleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLys 379  
 Db 706 ATCAGCTGGCAACCTTGCACCATCTCTTTCATCAGATCAACACCGGCTATCCAAAG 765  
 Qy 380 SerTrpIleValProGlyTyrPargLeuGlyTyrValAlaValTyrAspProArg----- 397  
 Db 765 GCATACCGGTGCGCAGGATACCGAGTGTGCTGGATGATGATGATGATGATGATGATGAT 825  
 Qy 398 -----LysIleLeuGlnGluThrLysIleSerThrSerIle 409  
 Db 826 GCAGGTGGATTATTGAGGGCTCGAAGCTCTCGAGGACCTCGACTCTGCCCAATGTC 885  
 Qy 410 ThrAsnTyrLeuAsnValSerThrAspProAlaThrPhe---IleGlnAlaLeu--- 427  
 Db 886 -----CCAGCTCAGCAGCGCTATTTCAGGTAGCTCTCGGT 918  
 Qy 428 -----ProGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeu 445  
 Db 919 GGAGCGCAGTCCATCTACGACCTCTGCGCAACAC-----GCCGA 960  
 Qy 446 LeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCys 465  
 Db 961 CTCTCTGAACAGCGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 Qy 466 ProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGlu 485  
 Db 1021 GTG---AAACCAATGGGAGCTCTATACGCGTTCCCAAGCTCGACCCCAAGTGTAC--- 1074  
 Qy 486 GluIleAspAspIleAspPheCysCysLysLeuAlaLysGluGluSerValIleLeu 505  
 Db 1075 GAAATCCAGCAGACACCAACTCATGCTGATGATGATGATGATGATGATGATGATGATGAT 1134  
 Qy 506 CysProGlySerValLeuGlyMetAlaAsnTrp-----ValArgIleThr 520  
 Db 1135 GTTCAGGGCAGCTGGCTTC-----AACTGGCCACATCAGCATCACTTCCGAGTGGTC 1185



QY 521 PheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGln 540  
Db 1186 ACCCTGCCATGCCATCCAGTTGGAAACGCAATTGAGCGCTGGGTAACTCTCTGTCC 1245  
QY 541 ArgAsnLysLys 544  
Db 1246 ACTTACAAGCAG 1257

RESULT 15  
US-10-156-761-1  
; Sequence 1, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; PRIOR FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 1  
; LENGTH: 9025608  
; TYPE: DNA  
; ORGANISM: Streptomyces avermitilis  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4187715)  
; OTHER INFORMATION: a, t, c, g, other or unknown  
US-10-156-761-1

Alignment Scores:  
Pred. No.: 1.69e-19 Length: 9025608  
Score: 329.00 Matches: 148  
Percent Similarity: 35.31% Conservative: 96  
Best Local Similarity: 21.42% Mismatches: 265  
Query Match: 11.52% Indels: 182  
DB: 14 Gaps: 23

US-10-019-783-2 (1-551) x US-10-156-761-1 (1-9025608)

QY 5 ArgGlnSerAspGlyValAla--AlaAsnGlyLeuAlaValAlaAlaAlaAsnGly 23  
Db 4086387 CGGAGCGGACGCGCTCACCGCGGAGCGCGCTCGCGCGTACCGCGCTCC----- 4086440  
QY 24 LysSerAsnGlyHisGlyValAlaAlaAlaValAsnGlyLysSerAsnGlyHisGly--- 42  
Db 4086441 ---GTCGAGGCGCGCGAAGTCTGTGGCGCACCGGTTCTTCTCCCGGAGGGAGAGAC 4086497  
QY 43 -----ValAspAlaAspAlaAsn--GlyLysSerAsnGlyHisGlyValAla 58  
Db 4086498 ACCGAGCGCGGAGAGAGACGCGTCTCGGACCGGCGGCGGTGACGGCGGGCGG 4086557  
QY 58 laAspAlaAsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAsnGlyHisGlyGlu- 77  
Db 4086558 CCGCGCTCCGACGCGCGGTGTGGCGCGCGCGCGCGCGCTCGCGTCCACGAGCGG 4086617  
QY 78 -----AlaThrAlaAsnG 82  
Db 4086618 TGCTCGGACCGCTCGGCGCTCGGCGCACCGCGCTCGCGCTCGACCCCGGAGCGGCGG 4086677  
QY 82 lYlYsThrAsnGlyHis---ArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnG 101  
Db 4086678 GACCCGAGCGGAGACCCCGCGATCGGACGCGCGCACCCAGGCGCGAGAC---CGCG 4086734

QY 101 lYgluSerAsnGluHis---AlaGluAspSerAlaAlaAsnGlyLysSerAsnGlyHisA 120  
Db 4086735 CCGAGCGGCGAGAGACCGCGGAGCGTGGCGCGTCCGCGACCCCTTCGGCGCGCGCGCGT 4086794  
QY 120 laAlaAlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspG 140  
Db 4086795 CGCGTCCGCGCTCCAGGACCGGCGCGCGCGCGCGCGCGTTCGTCGCGTCCGGAACG 4086854  
QY 140 lY----- 140  
Db 4086855 CTCACGGGCGAGCTCGCGGTCTGTCTAGGCGTTCTCCACCGGTGGCGCGCGCGCGG 4086914  
QY 141 ----ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIle-----ArgTyrL 157  
Db 4086915 TCCCGGTACTCTCTGTCAGCGGCGCGCGCGCGCGCGCGGTGTGTAGAGCTCGCGCGG 4086974  
QY 157 ysIleSerAlaSerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyA 177  
Db 4086975 CGGTGGCGGATCGGCACGAGCGCGCGCGCGCGCGCGTTCGCGCCATTGTTCGCGGC 4087034  
QY 177 spPro----- 178  
Db 4087035 GCCCGGTGGCGGCTGAAGTGCATGACATGCGCGGAAAGTGTAGAGACAAGAGGGG 4087094  
QY 179 --SerValPheProAlaPheArgThrAlaValGlu-----AlaGluAspA 193  
Db 4087095 GTCACGTGTACCGCGGTGGCGCACCGCTCGACACCCCGTGTCTCAGTGGCGGAAACA 4087154  
QY 193 laValAlaAlaAlaValArg----- 199  
Db 4087155 CGCGCGTAAAGGCGTTACGCGCGCGGTAAGCTCGCGGACATGCAGGTGATCCAGTCCACGA 4087214  
QY 199 ----- 199  
Db 4087215 AGCTGCCAACGTCGTGTACGAGATCGGGGCGCGGTGTCTGAGAGAGCGATCGCGTGG 4087274  
QY 200 -----ThrGlyGln-----PheA 204  
Db 4087275 AAGCGGTGTCTCATCGATCTCTCAAGTCAACACGCGCAATCGCGCGGCTTCGGTTCG 4087334  
QY 204 snCys-----TyrP 207  
Db 4087335 AGTGGCGCGCGGAGATCTCTCGAGGACATCTCTCGCAACGCTCTCTCCGACAGCGGTACG 4087394  
QY 207 roAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyV 227  
Db 4087395 GCGACGCGAAGGCGCTGTGGCGCGCGCGCGCGCGTCTCATGAC---AACGACGCC 4087451  
QY 227 alProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyThrGlnAlaIleG 247  
Db 4087452 TCGGCTATCGAGACGCGACGTCGACGCTCTCTGTGGCAACGCGTCTCCGAGCTCATCG 4087511  
QY 247 luValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProG 267  
Db 4087512 TGTATGCCATGCGGCGCTGTGGAGC---GGGACGAGGTGTCTGTACCGGCGCGCG 4087568  
QY 267 lYlYrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspL 287  
Db 4087569 ACTACCGGTGTGACCGCGCTCGCTCTCTCTCGCGCGCACCGCGCGTCACTACCACT 4087628  
QY 287 euIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysA 307  
Db 4087629 GCGACGAGCAGTCGCGACTGATGCGCGACCTCGCGAGCGTGTGAGCGCAAGTCAACCGAC 4087688  
QY 307 snThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrA 327  
Db 4087689 GCACCAAGCGCATCTCATCATCAACCCGACCAACCCGACCGCGCGGTGTATGACGAGG 4087748  
QY 327 sPHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspG 347  
Db 4087749 CGATCTCGCGCGCTGACGAGCATCGCGCGCGCGCACAACTGCTGCTCTCTCTCGGACG 4087808  
QY 347 luValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyH 367

Mon Nov 10 11:53:09 2003

```

Db 4087809 AGATCTACGACAGATCTCTTACGCGCGCCACGACACACCCCGCGGTGCGCC 4087868
QY 367 isileThrProValLeuSerileGlySerLeuSerLysTrpIleValProGlyTrpA 387
Db 4087869 CCGATCTGCTCAGCTCAGCTTCAACGGCAGTGCAGAGGGGTACCGGGTGGCGGCTACC 4087928
QY 387 rgLeuGlyTrpValAlaValTrpAspProArg----- 397
Db 4087929 GGGTCGGCTGGATCGCATCTCCGGCGCGCGCGCCACGCGACTCGTACATCGAGGGCC 4087988
QY 398 --LysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal---- 415
Db 4087989 TGACGATCTTGGCGGAACATGCGGTGTGCGGGAACATGCGGGGTACGACACGGGGTGTGCG 4088048
QY 416 -----SerThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnT 434
Db 4088049 CCGCGCTGAGCGCGCTCAGACGATCAACGAACCTGGTGTCTCCCGGC----- 4088095
QY 434 hrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerCluIleCyst 454
Db 4088096 -----GGACGGCTCAGGAACACGGGACGTGGCG 4088126
QY 454 yrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetP 474
Db 4088127 AGGACTGCTGACCCAGATCCCGCGGTGAGCTGTGTG---AACCCGAAGGGCGGCTCT 4088183
QY 474 heValMetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAspIleAspPheC 494
Db 4088184 ATCTCTTCCCGGGGTGACCCCAAGGTCTTC---AAGATCAAGGACGACCGGCGATGG 4088240
QY 494 ysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetA 514
Db 4088241 TGCTGACTGTGTCGCCCGGAGAGATCATGGTGTCCAGGGCACGGGCTTC----- 4088293
QY 514 laAsnTrp-----ValArgIleThrPheAlaCysValProSerSerLeuG 529
Db 4088294 --AAGTGGCGGACCCCGACCACTTCCGGGTGTGTGACCTGCCGACGCTGGCGGATCTGA 4088351
QY 529 lnAspGlyLeuGlyArgIleLysSerPhe 538
Db 4088352 CCTCGCGGATCACCCCGGATCGGGAACCTC 4088380

```

Search completed: November 9, 2003, 14:35:07  
Job time : 20140.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 20:02:05 ; Search time 93.6482 Seconds  
(without alignments)  
2596.975 Million cell updates/sec

Title: US-10-019-783-2

Perfect score: 2856

Sequence: 1 MATVROSDGVAANGLAFAAA.....LGRKSFQCRNKRNSDDC 551

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 569378 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10019783/runat\_07112003\_161644\_19870/app\_query.fasta\_1.1358  
-DB=Issued Patents NA -QFMT=fastap -SUPFIX=p2n.rni -MINMATCH=0.1 -LOOPCI=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10019783 @CGN1\_1\_76 @runat\_07112003\_161644\_19870 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	360.5	12.6	4403765	3	US-09-103-840A-2
C 2	360.5	12.6	4415229	3	US-09-103-840A-1
C 3	348	12.2	1185	2	US-08-646-590B-39
C 4	348	12.2	1185	3	US-09-412-184-39
C 5	330	11.6	1236	4	US-09-107-532A-1288
C 6	330	11.6	1482	4	US-09-328-352-2537
C 7	324	11.3	7818	4	US-09-634-238-14
C 8	317	11.1	1664976	4	US-08-916-421B-1
C 9	314.5	11.0	7939	4	US-08-961-527-9
C 10	314	11.0	1527	4	US-09-252-991A-9583
C 11	314	11.0	1677	4	US-09-252-991A-9544
C 12	295	10.3	981	4	US-09-252-991A-9707
C 1	360.5	12.6	4403765	3	US-09-103-840A-2
C 2	360.5	12.6	4415229	3	US-09-103-840A-1
C 3	348	12.2	1185	2	US-08-646-590B-39
C 4	348	12.2	1185	3	US-09-412-184-39
C 5	330	11.6	1236	4	US-09-107-532A-1288
C 6	330	11.6	1482	4	US-09-328-352-2537
C 7	324	11.3	7818	4	US-09-634-238-14
C 8	317	11.1	1664976	4	US-08-916-421B-1
C 9	314.5	11.0	7939	4	US-08-961-527-9
C 10	314	11.0	1527	4	US-09-252-991A-9583
C 11	314	11.0	1677	4	US-09-252-991A-9544
C 12	295	10.3	981	4	US-09-252-991A-9707

13	295	10.3	1173	4	US-09-107-532A-1579	Sequence 1579, Ap
14	289	10.1	1107	4	US-09-724-623-4	Sequence 4, Appli
15	289	10.1	1830121	4	US-09-557-884-1	Sequence 1, Appli
16	289	10.1	1830121	4	US-09-643-990A-1	Sequence 1, Appli
17	287.5	10.1	1122	1	US-08-599-171A-18	Sequence 18, Appl
18	287.5	10.1	1122	2	US-08-646-590B-18	Sequence 18, Appl
19	287.5	10.1	1122	3	US-09-068-226-18	Sequence 18, Appl
20	287.5	10.1	1122	3	US-08-412-184-18	Sequence 18, Appl
21	287.5	9.9	1170	4	US-09-724-623-5	Sequence 5, Appli
22	267	9.3	1032	4	US-09-107-532A-1892	Sequence 1892, Ap
23	265	9.3	1257	4	US-09-328-352-399	Sequence 399, App
24	264	9.2	1206	4	US-09-252-991A-10627	Sequence 10627, A
25	255.5	8.9	1197	1	US-08-599-171A-21	Sequence 21, Appl
26	255.5	8.9	1197	2	US-08-646-590B-21	Sequence 21, Appl
27	255.5	8.9	1197	3	US-09-068-226-21	Sequence 21, Appl
28	255.5	8.9	1197	3	US-09-412-184-21	Sequence 21, Appl
29	244	8.5	1149	4	US-09-107-532A-2417	Sequence 2417, Ap
30	233.5	8.2	1233	4	US-09-252-991A-16539	Sequence 16539, A
C 31	233.5	8.2	1770	4	US-09-252-991A-16060	Sequence 16060, A
C 32	233.5	8.2	28473	4	US-08-961-527-83	Sequence 83, Appl
C 33	229.5	8.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
34	228	8.0	1505	4	US-09-142-481-1	Sequence 4, Appli
35	226	7.9	1491	2	US-08-941-647A-4	Sequence 4, Appli
36	224	7.8	1975	3	US-09-109-204-4	Sequence 4, Appli
37	224	7.8	1975	4	US-09-490-032-4	Sequence 4, Appli
C 38	224	7.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
C 39	222.5	7.8	1197	4	US-09-252-991A-6207	Sequence 6207, Ap
40	222.5	7.8	1323	4	US-09-252-991A-6342	Sequence 6342, Ap
41	221.5	7.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
42	220.5	7.7	1701	3	US-08-599-968-2	Sequence 956, App
43	218	7.6	1164	4	US-09-328-352-956	Sequence 10807, A
44	212.5	7.4	1389	4	US-09-252-991A-10807	Sequence 1, Appli
45	212	7.4	1748	3	US-08-765-889C-1	

#### ALIGNMENTS

RESULT 1  
US-09-103-840A-2/C  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Alignment Scores:  
Pred. No.: 1.34e-21  
Score: 360.50  
Percent Similarity: 38.36%  
Best Local Similarity: 24.26%  
Query Match: 12.62%  
DB: 3  
Length: 4403765  
Matches: 148  
Conservative: 86  
Mismatch: 241  
Indels: 135  
Gaps: 21  
US-10-019-783-2 (1-551) x US-09-103-840A-2 (1-4403765)

18 AlaAlaAlaAlaAsn-GlyIysSerAsnGlyHisGly-----ValAlaAlaAlaAla 34  
 403781 GCCGACGGCGTCCAGGTCTGCACGTGGCCAGGTGCTACTCGGATCACTAGACCCAG 403722  
 34 LasnGlyIysSerAsnGlyHisGlyValAlaAspAlaAsnGlyIysSerAsnGlyHis 54  
 403721 ACAAGGCGCAGCTCCGCGCCAGGGTAC-----GGCGCCAAACAGGCTC 403677  
 54 sGlyValAlaAlaAlaAspAlaAsn-----GlyIysSerAsnGlyHisAlaGlu---AlaTh 71  
 403676 AGGAGCGCGCCCGAAAGCCGCTCTAAGGCTGCCGACCGCTGACCGCGCTGGAGGCAC 403617  
 71 rAlaAsnGlyHisGlyValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 91  
 403616 CCGCGAGGACCAACAGCGCGCAGCGCC-----GGCGCTCCCGCAGCACCCTGTAAGGCC 403560  
 91 nGlyHisAlaGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 109  
 403559 TCGGATCG 403500  
 109 p-----SerAlaAla 112  
 403499 CCGCGCTCG 403440  
 112 AsnGlyGlu-----SerAsnGlyHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 129  
 403439 CCGGCGGAAAGAACACACTCCCGCGCTCTCGGCTCGCGCAGCGCAGCGCGCGCGCGCG 403380  
 129 aValGluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAlaAlaAlaAlaAla 149  
 403379 AACCCGAAGCAAA-----GCCGCAACCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 403347  
 149 tSerIleAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 168  
 403346 CACCGCAAGCG 403287  
 168 ----- 168  
 403286 TGGGATTCG 403227  
 169 -----ProValLeuProLe 173  
 403226 AGCAATTCAGTGCACACAGATGACCAATTTGGAGCTGACCAACCCAGCTGCGCT 403167  
 173 uAlaHisGlyAspProSerValPhePro-AlaPheArgThrAlaValGluAlaGluAspA 193  
 403166 GGCACACCGCCAGCCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 403107  
 193 laVal-----AlaAlaAlaValArgThrGlyGlnP 203  
 403106 TCCTGTACAGATCCG 403047  
 203 heAsnCysThrProAlaGlyValGlyLeuProAla-----AlaArgSera 218  
 403046 ACCGATCTCTAACTACATCTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 402987  
 218 laValAlaGluHisSerGlnGlyValProThr----- 229  
 402986 TGATCATCGCGGATATCATCCAGCGCGCTGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCG 402927  
 230 --MetLeuSera----- 233  
 402926 GCATCTGTGCG 402867  
 234 -----AspAspValPheLeuThrAlaGlyThrGlnAlaAlaAlaAlaAlaAlaAla 402807  
 402866 GATTCGAGCTGGAGCGAGCTACTACCTGGTAACGGGCTCTCCGAGCTGATCAGATGACGC 402807  
 250 leProValLeuAlaGlnThrAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 270  
 402806 TGCACGCCCTGTGGCAAC-----GGCATCAGTGTCTGATTCATCAGCGGCTACCGCGC 402750  
 270 snTyrGluAlaArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 290

402749 TGTGACCGGTTCGACCTCGTGGCTGGCGCACTCGGTCACTACCTGTCGATGAGA 402690  
 290 sPlyGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspIysAsnThrAla 310  
 402689 CCCAAGGCTGGCAGCCGATATCGCGACCTGGATCCAGATCAGCGCGCCACCAAG 402630  
 310 laMetValIleIleAsnProAsnAsnProCysGlySerValTySerTyAspHisLeuS 330  
 402629 CCTGTGTGTGTATCAACCCACACACCCGCGCGGTGTACAGCTGCGAAATCTCTCA 402570  
 330 exLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyG 350  
 402569 CCAGATGTGTGATGTGCGCGCGCAAGCATCACTGCTGTGGCGGACCAATCTACG 402510  
 350 lyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrP 370  
 402509 ACAAAATCTCTTACGACGACGCAAGCAGCATCAGCTGGCATCGTCCCGGATATGT 402450  
 370 roValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyT 390  
 402449 TGTGCTGACCTTCAATGGCTGTCCAGGCTACCGGTGCGCGGATACCGGCGCGCT 402390  
 390 rpValAlaValTyAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleT 410  
 402389 GGTGCGCATCACCGACCCCAAGGAG-----CACGCCAGAGCTTCATCGAGGCGATCG 402336  
 410 hrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnI 430  
 402335 GCTGTGTCCTATGCGGTGTGCCAAATGTCCGCGCCAGCATGCCATT---CAGG 402279  
 430 leLeuGluAsnThrLysGluAspPhePheLysAlaIleIle-----GlyLeuLeuLysG 448  
 402278 TTGCACTGGCGCGCCATCATCAGACATCGAGGACCTGCTGCCCGCGCGCGCTGCTCG 402219  
 448 luserSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisL 468  
 402218 AGCAGCGGACATCGCTCGGACCAAGCTCAAGAGATCCCGGGGTGCTGCTCGTC---A 402162  
 468 ysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluIleAla 488  
 402161 AACCGCGCGCGCTGTATGCTTCCCGGTAGACCCCGAGGTCTAC---GACATCG 402105  
 488 spAspAspIleAspPheCysCysLysLeuAlaLysGluSerValIleLeuCysProG 508  
 402104 ACACGACGAGCAACTCGTGTGATCTGCTGTCGAGAGATCTGTCGAGAGATCTGTCACCCAG 402045  
 508 lySerValLeuGlyMetAlaAsnTrp 516  
 402044 GCACCGGGTTC-----AACTGG 402028

RESULT 2  
 US-09-103-840A-1/c  
 ; Sequence 1, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R. M.  
 ; APPLICANT: FRASER, John C.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; FILE REFERENCE: 24366-20007.00  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: Patent ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 4411529  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; OTHER INFORMATION: H37Rv  
 US-09-103-840A-1



STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: Windows95  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,590B  
FILING DATE: 08-May-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/599,171  
FILING DATE: 09-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/01094  
FILING DATE: 21-January-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Ph.D., Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 09010/017001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1185 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...1182  
US-08-646-590B-39

Alignment Scores:  
Pred. No.: 3,888-26 Length: 1185  
Score: 348.00 Matches: 105  
Percent Similarity: 47.72% Conservative: 83  
Best Local Similarity: 26.65% Mismatches: 166  
Query Match: 12.18% Indels: 40  
DB: 2 Gaps: 15

US-10-019-783-2 (1-551) x US-08-646-590B-39 (1-1185)

QY 158 llsAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAla 174  
Db 58 ATAACCGCAAAAGCAAGAAATTAAGGGCTAAAGGAGTG---GACGTTATAGGTTTGA 114  
QY 175 HsGlyAspProSerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaVal 194  
Db 115 GCGGGAACCTGAC-----TTGCACACCCGACTTCATAAGGACCTGT 162  
QY 195 AlaAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAla 214  
Db 163 ATAAGGCTTTAAGGAAGAAAGCAAG---TACGCTCCCTCCGCGGAATACAGAG 219  
QY 215 AlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAsp 234  
Db 220 CTCAGAGAGCTATAGCTGAAACCTACGAAAGAAACAAAGTTGAGTACAACTTCA 279  
QY 235 AspValPheLeuThrAlaGlyGlyThrGlnAlaAlaGluValLeuLeuProValLeuAla 254  
Db 280 GAGATAGCTTTCCGACGAGCGAAATGTTCTCTTCTCTCATATTCATGCTATAC 339  
QY 255 GlnThrAlaGlyAlaAlaHisLeuLeuProArgProGlyTyrProAsnTyrGluAlaArg 274  
Db 340 GAC---GAAGGAGACGAGGTTTATACCTACGCTTACTGGGTAACTTACCCGAA--- 393

QY 275 AlaAlaPheAsnArgLeuGluValArgHisPheAspLeuLeuPro----- 289  
Db 394 -----CAGATAAGTTCTTCGAGGGGTTCCGTTGAGGTTCCTCTCTA 435  
QY 290 -----AspLysGlyTyrGluLeuAspLysSerLeuGluSerLeuAlaAspLysAsn 307  
Db 436 AAGAAAGAGAAAGGATTTCAATTAAGTCTGGAAGATGTGAAAGAAAGGTTTACGAGAGA 495  
QY 308 ThrThrAlaMetValLeuLeuAsnProAsnProCysGlySerValTyrSerTyrAsp 327  
Db 496 ACAAAAGCTATAGTCTAACTCCGAAACACCCACTGCTGCTGTTTACGAGAGAGAG 555  
QY 328 HisLeuSerLysValAlaGluValAlaLysArgLeuGlyLeuValLeuAlaAspGlu 347  
Db 556 GAACITTAAGAAATAGCGAGTTTTCGCGTGGAGAGGGCATTTTCATAATTTCCGATGAG 615  
QY 348 ValTyrGlyLeuValLeuGlySerAlaProPheLeuProMetGlyValPheGly--- 366  
Db 616 TGTATAGTACTTCGTTTACGCTGATGCAAAATTTGTAGCCCTGCTCTTTCGGAT 675  
QY 367 -----HisIleThrProValLeuSerIleGlySerLeuSerLysSerTrpIleVal 383  
Db 676 GAAGTAAGAACATACAC-----TTCACGGTAAACGCTTTTTCGAGAGCTATTCATG 729  
QY 384 ProGlyTyrArgLeuGlyTyrPValAlaValTyrAspProArgLysIleLeuGlnLthr 403  
Db 730 ACTGGTTGGCAATAGGTTATAGCTGCCCGAA-----GAGTACGCA 774  
QY 404 LysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIle 423  
Db 775 AAGTATAGCAGAGCTT---AACAGCCAGAGTGTTCC---AACGTCACACTTTCGCC 828  
QY 424 GlnAlaAlaLeuProGlnLeuGluAsnThrLys---GluAspPhePheLysAlaIle 442  
Db 829 CAGTATGGAGCTCTTGAGGCTTGAAATAATCCAAAGTCTAAAGATTTTGTAAACGAAATG 888  
QY 443 IleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyr 462  
Db 889 AGAATGCTTTTGAAGAGGAGAGGATACGGCTGTAGAGAGCTTCTTAAATTCAGGT 948  
QY 463 IleThrCysProHisLysProGlySerMetPheValMetValLysLeuAsnLeuHis 482  
Db 949 ATGGATGTGTA---AAACCCGAGAGTGCTTTTACATATTTCCGAGCTTCTCCGCTTAC 1005  
QY 483 LeuLeuGluGluLeuAspAspLysPheCysCysLysLeuAlaLysGluGluSer 502  
Db 1006 ---GCTGAGAAACTGGGTGCTGATGTGAAACTCTCGGAGTCTCTCTGAAAGGGCTAAG 1062  
QY 503 ValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIleThrPheAla 522  
Db 1063 GTTCCGGTGTCCCGTTCCGGCTTCGAGCTCCCGGATTTTCAGGCTTCTTACGCG 1122  
QY 523 CysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536  
Db 1123 CTTTCCGAGAAAGACTCGTTGAGGTTATAGAGAAATAAG 1164

RESULT 4  
US-09-412-184-39  
; Sequence 39, Application US/09412184  
; Patent No. 6268188  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Patrick V.  
; APPLICANT: Swanson, Ronald V.  
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037



OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107.532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 1288:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1236 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...1236  
SEQUENCE DESCRIPTION: SEQ ID NO: 1288:  
US-09-107-532A-1288  
  
Alignment Scores:  
Pred. No.: 2,93e-24 Length: 1236  
Score: 330.00 Matches: 108  
Percent Similarity: 46.08% Conservative: 92  
Best Local Similarity: 24.88% Mismatches: 181  
Query Match: 11.55% Indels: 53  
DB: 4 Gaps: 17  
  
US-10-019-783-2 (1-551) x US-09-107-532A-1288 (1-1236)  
QY 131 GIUTPAsnPheAlaGlyValLeuAlaAlaThrGlyAlaAsnMetSer 150  
DB 39 GAATGG-----GATGAGTTATCAAA-----AAGACGAAACG 71  
QY 151 IleArgAlaIleArgTyr-LysIleSerAla-----SerValGlnGluLysGl 166  
DB 72 ATTGAACCATCTGTTACTTTGGCAGCTCGCGCAAAAGCCGAGGCTTTAAAGCAAAAGG 131  
QY 166 yProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgTh 186  
DB 132 A---AGAGATGTCGTGAGTTTACTGTGGGAGAACCGGAT-----TTTGGCAC 176  
QY 186 yAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTy 206  
DB 177 ACCAGAAACATCCCAAGACGACGAATCGAAGCAATCCGTAACGGAAAGTAGTATTA 236  
QY 206 rProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGl 226  
DB 237 TACACCAACAGCAGGAATCCCAAGACTTCGACGAGCGGATCGTTGATTATCTCAGAAAAA 296  
QY 226 yValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIle 246  
DB 297 CGATGACCTTGCTATGATCTTCACAAACGGTCGTACAGATGCGCGAAATTTGCTTT 356  
QY 246 eGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgPr 266  
DB 357 GTATACACTTTTCAACAACATCTCGACCCGCAAGAT---GAAGTCATCATCTCTGTCGCC 413  
QY 266 oGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAs 286

Db 414 TTATTGGGTCAAGTTACGGCGAG-----CAAGTGAACCTGGCAGA 452  
QY 286 pleuilePro-----AspLysGlyTrpGluIleAspIleAspSe 299  
Db 453 AGGCCTGCTGCTTTGTGAAAGGAAAGAAATTAATAACTGGAAGTCACTGTGTAACA 512  
QY 299 rLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnPr 319  
Db 513 GTTGAACAACGACGCTTCAGGAAAAACAAAGCGGTGATCATCAATTCGCCATCGAATCC 572  
QY 319 oCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLe 339  
Db 573 GACTCGAATGATTTATAGCAAGATGAACATCGAGCAATTCGTAATGGGCGAGTAAAAA 632  
QY 339 uGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPh 359  
Db 633 TGATATTTTGTAGTTGCTGATGATATTATGTCGTTTATGTTATATGCAATGAGTT 692  
QY 359 eIleProMetGlyValPheGlyHisIle-----ThrProValLeuSerIleGlySerLe 377  
Db 693 TACTCAATTCACGATTCAGAGCAATCAAAAATCAGACCATCATCAATAATGGGT 752  
QY 377 uSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProAr 397  
Db 753 GTCTAAGACCTATGCGATGACAGGCTGCGCATCGGTTAT---GCGTTGGGAATCTCGA 809  
QY 397 glyIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerTh 417  
Db 810 AATTATCAATGGGATGATCGCTGTTGCTTCCCAATCACT-----AG 851  
QY 417 rAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGluAs 437  
Db 852 CAATCGACCGCAGTCAGTCAATATGACGCTGTTGAAGCATTTGAAAGGT---GAGCAAGA 908  
QY 437 pPhePheLysAlaIleIleGlyLeuLysGluSerSerGluIleCysTyrLysGlnIle 457  
Db 909 TACAGTAGAAGAAATCGGGAACGCTTTGAAGAACGCTGATATCGTTATATCCGCTGT 968  
QY 457 eLysGluAsnLysTyrIleThrCysProHisLysProGluGlySer-----MetPheVa 475  
Db 969 GGCAGAACTACCAAGGTGTTTCTCTT---GAAAAGCCGCAAGGAGCTTTTATCTTTTCC 1025  
QY 475 lMetValLysLeuAsnLeuHisLeu-----LeuGluGluIleAspAspIleAspPh 493  
Db 1026 TAATGTAAAGAAACACGCTTCGCATGTGTAAGTATGAAATGTACACAAATGGGTAGAA-- 1083  
QY 493 eCysCysLysLeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMe 513  
Db 1084 -----GATCTTTAGAAAGAACAGGGGTAGCACTTGTGACGCGGAAGGATTTGGGCG 1136  
QY 513 tAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGl 533  
Db 1137 ACCAGAAATGTCGCGATGAGTTATGCGACTGATCGGCTAGCTAGAGGACGCGGTACG 1196  
QY 533 yArgIleLysSerPheCysGluArgAsnLysLysArgAsn 546  
Db 1197 CAGATCAAGCAATTTATTGAG---AGTAAAGTCAAAAT 1233  
  
RESULT 6  
US-09-328-352-2537  
Sequence 2537, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 2537  
LENGTH: 1482



TYPE: DNA  
 ORGANISM: Acinetobacter baumannii  
 US-09-328-352-2537

Alignment Scores:  
 Pred. No.: 3,94e-24 Length: 1482  
 Score: 330.00 Matches: 104  
 Percent Similarity: 45.43% Conservative: 80  
 Best Local Similarity: 25.68% Mismatches: 181  
 Query Match: 11.55% Indels: 40  
 DB: 4 Gaps: 12

US-10-019-783-2 (1-551) x US-09-328-352-2537 (1-1482)

QY	151	IleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArgProVal	170
Db	319	ATTGCGGACACGATTATACAGCGCGCAATGAATGGAAGCAAGGACATAAA---ATC	375
QY	171	LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla	190
Db	376	ATTAAGCTGAATATCGCAACCTGCTCCATT-----GGTTTGAAGCA	420
QY	191	GluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCys-----TyrProAla	208
Db	421	CCACAAGAAATTTAATGACGTTGCTTTAAACCTACCAAAATGCAATGTTATGATGAT	480
QY	209	GlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValPro	228
Db	481	TCAAAAGTATCTTTCCGCGACGTAAGCAATCTGCCGATGATACACGAAAAAGGCATC	540
QY	229	TyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluVal	248
Db	541	TTAAATATGACGTTAATGACGTGATATCGGTAATGGTGTATCTGAACTCATTTGTCATG	600
QY	249	IleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyr	268
Db	601	GCTATGCAAGCCATCTCGATGAT---GGTGATGAATGCTCATTCGATCCCGGACTAT	657
QY	269	ProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIle	288
Db	658	CCGCTTTGGACAGCTCCGTTAACTGTACGCGGTACAGCAATTCATTATAAGTGAT	717
QY	289	ProAspLysGlyTrpGluIleAspLeuSerLeuGluSerIleAlaAspLysAsnThr	308
Db	718	GAAAGAAACAGTTGGTATCCTGATATGCCGATATTGAAAGTAAATTTACTTCAACACT	777
QY	309	ThrAlaMetValIleIleAsnProAsnProCysGlySerValTyrSerTyrAspHis	328
Db	778	CGTGGTATTGTCATTATCAATCCGAAACCAACCACTGTTGGTATATCCAGTCATGTG	837
QY	329	LeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluVal	348
Db	838	CTTGACAAATTTGGSCACTTCCAAAGAAACATGACCTGATTTATTGCTGACGAAAT	897
QY	349	TyrGlyLysValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIle	368
Db	898	TACGACAAATCGTTTACATGGCATTGAACATGTGTGTGTCTTTAGCAGGCGAC	957
QY	369	ThrProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeu	388
Db	958	CAATTATGTAATTTCTTCAATGGGCTATCAAAAGCATATCGTATTGCTGGTTCCGTTCA	1017
QY	389	GlyTrpValAlaValTyrAspProArgLys-----IleLeuGlnGluThrLys	404
Db	1018	GGCTGATGGCAATCACAGCGATGAAGCGGTGCAGCAGACTATATCGAAGGTCTGGAT	1077
QY	405	IleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGln	424
Db	1078	ATGTTAGCTCAATCGGTTTATGTGCAACCGTACAAGCTCAGTATGCA-----ATTCAA	1131
QY	425	AlaAlaLeuProGlnIleLeuGlnAsnThrLysGluAspPhePheLysAlaIleGly	444
Db	1132	ACGGCGCTT-----GGTGGCTATCATGCTTCTATTAATGAT	1164

Db 6790 AAGGT---GCCGACGTCATTGATTGTCATTGGCCCAACGAT-----TTT 6746  
Qy 185 ArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsn 204  
Db 6745 TCAACCCCTAAGCGGATGATGACGCGATTCGCGCGATTCAGCTGCGTGAATGCGAGT 6686  
Qy 205 CysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeu--- 223  
Db 6685 TTCATACGCGACGACCGGTAATCCGGAATTAAGCAGCGGATTAAGTACCGGATATT 6626  
Qy 224 ---SerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 242  
Db 6625 GCCCAAGACGCTATTCGTTAT-----GATCATCGTCAATTCGTCACACCGCGCT 6572  
Qy 243 ThrGlnAlaIleGluValIleIleProValIleAlaGlnThrAlaGlyAlaAsnIleLeu 262  
Db 6571 AAGTTTGTTCGTTATGCTTATTCAGGTTTCTTA---AACCCAGCGGATGAGTGTG 6515  
Qy 263 LeuProArgProGlyTyrProAsnTyrGluAlaArgAlaPheAsnArgLeuGluVal 282  
Db 6514 ATTCCTGTCCATACGCTGGGTTCTACGAGCAACAGATTAAATTG---GGGACGCGGTG 6458  
Qy 283 ArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSer 302  
Db 6457 CCACATCTG---GTCATGCCGCGAGTCGACATTAAGTCAGTTCGATGATCTTGAGCG 6401  
Qy 303 IleAlaAspLysAsnThrAlaMetValIleIleAsnProAsnAsnProCysGlySer 322  
Db 6400 GCTCGGACCGATAAACCCGGCATGATTAATCAATTCGCCACAAAACCAAGTGGCGT 6341  
Qy 323 ValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeu 342  
Db 6340 GTCTATGATCGCACGGAACGCTTAAATGGCAATGGCGCTGAAGCATCATATTTG 6281  
Qy 343 ValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMet 362  
Db 6280 GTAGTGACTGACGATATTTACGAGATCTGATTTATTAACGGTACGACTTACACCTCAATG 6221  
Qy 363 GlyValPheGlyHisIleThrProValLeuSer-----IleGlySerLeu 377  
Db 6220 -----ATTAGTATCGATCCCGATATCCGACGGAATACGTGTTTAATTCGCGCGTC 6170  
Qy 378 SerLysSerTrpIleValProGlyTyrTrpArgLeuGlyTrpValAlaValTyrAspProArg 397  
Db 6169 TCCAGTATATGCGATGACCGGTGGCGGATGGTGGTATGCGGC-----GGTCCGAA 6116  
Qy 398 LysIleLeuGln-----GluThrLysIleSerThrSerIleThrAsnTyrLeuAsnVal 415  
Db 6115 AAGCTGATTCAGGCCATCGCGACCTTTATTAGCCACACGACCTCTAATCCGCGACGAGT 6056  
Qy 416 SerThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLys 435  
Db 6055 TCCGAATACGCGCGGTGGCGAGCTTTAACTCGGATCAGCAGGTGTGTGAAGAATCGCG 5996  
Qy 436 GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLys 455  
Db 5995 CGTGCTTT-----GAAGAAGCGCTCAATCTTCTTCTATGAT 5960  
Qy 456 GlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheVal 475  
Db 5959 CTTCTGGCAGATATCCCGGTTTCGATATGGAGATAAACCGAGCGGCTTCTATCTT 5900  
Qy 476 MetValLysLeu-----AsnLeuHisLeuLeuGluGluIleAspAspIle 491  
Db 5899 TTCGCCAATATAAGCGTCGCGCTCAATTGATGATTTATGTCAGGTGAT----- 5849  
Qy 492 AspPheCysCysLysLeuAlaLysGlnGluSerValIleLeuCysProGlySerVal-Le 511  
Db 5848 GATTTTATCAGTGCACCTGTGACCGAAACCGGGTTGCCATTGTTCTCGGACGCGGTT 5789  
Qy 511 uGlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspG 531  
Db 5788 GGCATGCGCGATCATCGCGGATTAAGTATTGTAAAGATTGGCCACGCTCTGAAGAGCG 5729

Qy 531 yLeuGlyArgIleLysSerPheCysGlnArgAsn 542  
Db 5728 CGCCCGGCGTATCCGGAGTTTGTGTAATAAT 5695  
RESULT 8  
US-08-916-421B-1  
; Sequence 1, Application US/08916421B  
; Patent No. 6503729  
; GENERAL INFORMATION:  
; APPLICANT: Bult et al.  
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus  
; TITLE OF INVENTION: jannaschii  
; PATENT NO. 6503729  
; FILE REFERENCE: PB275  
; CURRENT APPLICATION NUMBER: US/08/916.421B  
; CURRENT FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 60/024,428  
; PRIOR FILING DATE: 1996-08-22  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1664976  
; TYPE: DNA  
; ORGANISM: Methanococcus jannaschii  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (28222)..(28222)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (28257)..(28258)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84773)..(84773)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84808)..(84808)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84812)..(84812)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98120)..(98120)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98159)..(98159)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98239)..(98239)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98266)..(98266)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98343)..(98343)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (103998)..(103998)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (148948)..(148948)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (163385)..(163385)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (191989)..(191989)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (191995)..(191995)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (231980)..(231980)

```

; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1095846)..(1095846)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1569020)..(1569020)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1602912)..(1602912)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664855)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-916-421B-1

```

## Alignment Scores:

Pred. No.:	8.05e-18	Length:	1664976
Score:	317.00	Matches:	91
Percent Similarity:	46.98%	Conservative:	80
Best Local Similarity:	25.00%	Mismatches:	159
Query Match:	11.10%	Indels:	34
DB:	4	Gaps:	13

US-10-019-783-2 (1-551) x US-08-916-421B-1 (1-1664976)

```

QY 182 ProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaValAlaValArgThrGly 201
Db 1469114 CCAGATTTCACACTCCAAAGCATATCATTTAGGCTGCAAAAGGCTTTAGATGAGGGG 1469173
QY 202 GlnPheAsnCysTyProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu 221
Db 1469174 AAAACTCAC---TACTCTCCAAACATGGAATTCAGACTTAGAGAGAGATAGCAAT 1469230
QY 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241
Db 1469231 AAGTTAAAGGATGATTACAACTTAGACGTTGATAGGACATATATTATTCTTACCTGTGA 1469290
QY 242 GlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261
Db 1469291 GCTTCAGAGCGCTTAATGTTATCTATTATGATCTTGTGTTGAC---AGAGGGATGAGGTT 1469347
QY 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281
Db 1469348 TTAATCCAAATCCATCTTTTGTGCTTATTTTCACTAACACAGCTTTGCTGAGGTTAAG 1469407
QY 282 ValArgHisPheAspLeuIleProAspIysGlyTrpGluIleAspIleAspSerLeuGlu 301
Db 1469408 ATTAAAGATATAGATTTA-----GATGAAACTTTAATATTGATTATAGAGAACTTAA 1469461
QY 302 SerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGly 321

```

Db 1469462 GAATCAATACTAAATAAACAAGCACTAATATATATTTAACTCTCAATCAAAATCTACTGGA 1469521  
Qy 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341  
Db 1469522 AAAGCTCATGATAAGGAGACCAATAAAGGCTGGCAGAGATTGCTGAGGATTATAATCTA 1469581  
Qy 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361  
Db 1469582 ATTATTTGTTTCAGATGAAGCTATATGAAGATTATCTACGATGAAGAGCATTACTCTCCA 1469641  
Qy 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTrp 381  
Db 1469642 ATGCAATTTACTGATAGA-----TGCAATATTAATTAACGGTTTCTCTAAAACCTAT 1469692  
Qy 382 IleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGln 401  
Db 1469693 GCCATGACTGGATGGAGATTGGATATTGGCTGTTTCTGAT-----GAA 1469737  
Qy 402 GluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer-----Thr 417  
Db 1469738 CTAATAAGGAGTAGATTATTAATCAACATATGATTAAAGATTCACTAGTATAGCTTTGCA 1469797  
Qy 418 AspProAlaThrPheIleGln---AlaAlaLeuProGlnIleLeuGluAsnThrLys--- 435  
Db 1469798 TGTGTACACACTTGTCTCAATATATGGGCGCATTTAGAGGTAGTCAAAAGTGT 1469857  
Qy 436 ---GluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyr 454  
Db 1469858 GTTGAGGATATGGTTAGAGAG-----TTTAAATGAGGAGAGATTAAATCTAT 1469905  
Qy 455 LysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPhe 474  
Db 1469906 AATGCTATTGAAGGAT-----ATCTTTAAAGTCAATTAAGCCAGATGGCGCATTTCTAT 1469956  
Qy 475 ValMetValLysLeuAsnLeuHisLeuLeuGluLeuLeuAspAspIleAspPheCys 494  
Db 1469957 ATATCCAGATGTTTCT-----GAGTATGAGATGGGTAGAGGTAGCT 1470001  
Qy 495 CysLysLeuAlaLysGluLysSerValIleLeuLysProGlySerValLeuGlyMet--- 513  
Db 1470002 AAGAAATTGAT--GAGAAATAGGTTTATGTTCTCTGAGTTGCAATTTGGTGAAT 1470058  
Qy 514 ---AlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeu 532  
Db 1470059 GGAGCTAACTATATTAGTTAGCTGCTACNAATATGAGATATAGAGAAAGCTTTA 1470118  
Qy 533 GlyArgIleLys 536  
Db 1470119 GGGATTATAAAA 1470130

## RESULT 9

US-08-961-527-9/c  
; Sequence 9, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527

; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7939 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-961-527-9  
Alignment Scores:  
Pred. No.: 2,38e-21 Length: 7939  
Score: 314.50 Matches: 102  
Percent Similarity: 45.50% Conservative: 90  
Best Local Similarity: 24.17% Mismatches: 182  
Query Match: 11.01% Indels: 49  
DB: 4 Gaps: 11  
US-10-019-783-2 (1-551) x US-08-961-527-9 (1-7939)  
Qy 125 GluGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAla 144  
Db 7225 GAAATGGAGAAAGTGTCACTTTG---GCTGCTGGAGCCCGTCCCAAGCACTGAAGCT 7169  
Qy 145 ThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGlu 164  
Db 7168 GAGGCT----- 7163  
Qy 165 LysGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 184  
Db 7162 ---AGAGATATTGTTCTTAACCTGGTGGTGAGCCAGAT-----TTC 7124  
Qy 185 ArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsn 204  
Db 7123 ACTATCTCCAAAATATCAAGATCCGCCATTGATCGATTCGAGATGGCGTCTTCT 7064  
Qy 205 CysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSer 224  
Db 7063 TTTTATACAGTAACCTCAGGCTGCGCAGAGCTTAAGCGCGCGGTCAATAGCTACTTGA 7004  
Qy 225 GlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGln 244  
Db 7003 CGCTTTTAGGGTATTCTGTAGCGTCAATCAAGTCAGACGCTGCGCGGAGCCAAATAT 6944  
Qy 245 AlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPro 264  
Db 6943 TCTCTCTATACCTCTTTATGCTGTGTGTC---AATCCAGGTGATGAAGTCATCATCCA 6887  
Qy 265 ArgProGlyTyrProAsnTyr---GluAlaArgAlaAlaPheAsnArgLeuGluValArgHi 284  
Db 6886 ACCCATACTGGGTAGCTATGAGATCAGGTCAAGATGAGGAGAGGGTCCCGTCTTT 6827  
Qy 284 sPheAspLeuIleProAspLysGlyTyrIleAspIleAspSerLeuGluSerIleAl 304  
Db 6826 GTTCTGTCTAAG-GAAGACAATCACTTTAAGGTGACCGTAGAGCAGTTAGAGACGCTCG 6768  
Qy 304 aspLysAsnThrThrAlaMetValIleLeuAsnProAsnAsnProCysGlySerValTy 324  
Db 6767 CACTGACAAGACCAAGGTTTGTGCTGCTGAATTCGCCATCTAATCCACAGGTATGATT 6708  
Qy 324 rSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValI 344  
Db 6707 CACCGGTGAGGAACTCTTGGCAATTGGAACTGGGCTGTAGAAAATGATATCTCATCT 6648

Qy	344	eAlaaspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVa	364
Db	6647	AGCAGCATATCTATGCGCGCTGGTTTATAATGGTCATGAGTTCACCGATTTCTAG	6588
Qy	364	lPheGlyHisIle-----ThrProValLeuSerIleGlySerLeuSerIlySerSsrTrpI	382
Db	6587	CCTATCGGAAGCGATTTCGAAGCAACACAGTGGTCATCAATGGTGTCTGCTAAACATTATGC	6528
Qy	382	eValProGlyTrpArgLeuGlyTTrpValAlaValTyrAspProArgIlySlyIleLeuGlnI	402
Db	6527	CATGACTGGTGGCGGATGGTTAT---GCCGTTGGAGAACGACACATTTTCGTGCGCAT	6471
Qy	402	uThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPh	422
Db	6470	GTCCAAGATTGCAGTCAAAACACT-----TCGAATCGTCAGCAGT	6429
Qy	422	eIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaI	442
Db	6428	AGCCCATATGCAGCAGTTGAGGCTCTATCAGGT---GAGCAAGATACTCTAGAAAGCAT	6372
Qy	442	eIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysT	462
Db	6371	CGGTCAGCGCCTTCAGGAACGCTTTAATACCATCTATCCCTCCCTTCAGAG-----	6320
Qy	462	rIleThrCysProHis-----LysProGluGlySerMetPheValMetVally	478
Db	6319	-----GTACCAGGATTGAAGTGGTCAACCCGACGGGGCCCTTCTATCTTCCCAA	6267
Qy	478	sLeuAsnLeuHisLeu---LeuGluGluIleAspAspIleAspPheCysCysLysLe	497
Db	6266	TGTCAAAAGGCCATGGAGATGAAGGCTACACGATGTGCAGACCTTACACTGTTAT	6207
Qy	497	uAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpVa	517
Db	6206	CTTAGAAGAAGCCGAAGTGGCTGGTTAAACAGAGAGCTGGATTTTGGACGACCAAGAAATGT	6147
Qy	517	lArgIleThrPheAlaCysValProSerSerLeuGlnAspClyLeuGlyArgIleLysSe	537
Db	6146	CCGCTCAGTATCGCAGACACGTAGACACGCTTAAAGAGCAGCTGCAACGCTTGAAGC	6087
Qy	537	rPhe	538
Db	6086	ATT	6083

RESULT 10

```

US-09-252-991A-9593
; Sequence 9583, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A

```

Alignment Scores:	
Pred. No.:	1.82e-22
Score:	314.00
Length:	1527
Percent Similarity:	43.58%
Best Local Similarity:	77
Query Match:	10.9%
Mismatches:	177
Indels:	56
Matches:	103
Conservative:	77

DB: 4 Gaps: 11

US-10-019-783-2 (1-551) x US-09-252-991A-9583 (1-1527)

Qy 151 IleArgAlaIleAArgTyLysIleSerAlaSerValGlnGlnLysGlyProArgProVal 170  
 |||||  
 Db 361 ATTGCGGGCGGTGCTCAGCAGCCAAACCGCTAGAGAGAGGTTCACCGC--ATC 417  
 |||||  
 Qy 171 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 190  
 |||||  
 Db 418 CTCAAGCTGAATATCGGCAACCCGCGCGCTC--GGTTCGAAGCTCCGAGGAAATT 474  
 |||||  
 Qy 191 GluAspAlaValAlaAlaValAArgThrGlyGlnPheAsnCysTyProAlaGlyVal 210  
 |||||  
 Db 475 CTCAGAGAGTTCATCCGCAACCTCCCGACCGCCAG-----GGCTACAGCACTCAAG 528  
 |||||  
 Qy 211 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyMet 230  
 |||||  
 Db 529 GGCCTGTTCAGCGCGCAAGGCAGTGCAGTATTCACGACGAGAGCAGGTGGAAGGC 588  
 |||||  
 Qy 231 LeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 250  
 |||||  
 Db 589 GTTGGCATCAGGACATCTACCTCGCAACCGGGGTGTCGGAACGTATCGTATGTCGCATG 648  
 |||||  
 Qy 251 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyProAsn 270  
 |||||  
 Db 649 CAGCGCTGCTGAACAAC--GGCGACGAGGTGCTGATCCCGCTCCGACTACCGCGTG 705  
 |||||  
 Qy 271 TyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProAsp 290  
 |||||  
 Db 706 TGGACCCCGCGGTGACCTCGCGCGGCAAGCGGTGCACTACCTGTGCGACGAGCAG 765  
 |||||  
 Qy 291 LysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThrAla 310  
 |||||  
 Db 766 GCCAACTGTGGCGGACCTGGAGACATCAAGGCGAAGATCAGCGCAACACACGAGCGG 825  
 |||||  
 Qy 311 MetValIleIleAsnProAsnAsnProCysGlySerValTySerTyAspHisLeuSer 330  
 |||||  
 Db 826 ATGTGTATCATCAACCGCAACACCCACCGCGCGGTGTATTCCAGGGAAGTGTGGAA 885  
 |||||  
 Qy 331 LysValAlaGluValAlaLysArgLeuGlyLysLeuValIleAlaAspGluValTyGly 350  
 |||||  
 Db 886 GGCATGTGAATCGGCCCGCCAGCAACCTGGTGTGTCTCCGACGAGATCTACGAC 945  
 |||||  
 Qy 351 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370  
 |||||  
 Db 946 AAGATCCTACAGCGCGCGCTCCACGTCTCCACCGCTCGTGGCGCGGACGTCTC 1005  
 |||||  
 Qy 371 ValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrp 390  
 |||||  
 Db 1006 TGCGTGACCTCAACCGCCCTGTCCAATCTACCGGGTGGCGCGCTCCGCTCCGGCTGG 1065  
 |||||  
 Qy 391 ValAlaValTyRAspProArgLysIleLeuGln----- 401  
 |||||  
 Db 1066 GTGGCATCTCCGGGCCCAAGCAGCGGCACAGAGCTATCGAAGGTCTCGACATCTC 1125  
 |||||  
 Qy 402 -----GluThrLysIleSerThrSerIleThr 410  
 |||||  
 Db 1126 GCCAACATGCGCTGTGCGCAACAGTCCCGCGCAGCAGCGCATCCAGACCGCCCTGGGC 1185  
 |||||  
 Qy 411 AsnTyRLeuAsnValSerThrAspProAlaThrPheIleGlnAlaLeuProGlnIle 430  
 |||||  
 Db 1186 GGCTACAGAGCATCAAC-----GATCTGGTCTTCGCCCGCG-- 1221  
 |||||  
 Qy 431 LeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSer 450  
 |||||  
 Db 1222 -----GGGCGCTCTGTGGAGCAGCGC 1242  
 |||||  
 Qy 451 GluIleCysTyRLeysGlnIleLysGluAsnLysTyRileThrCysProHisLysProGlu 470  
 |||||  
 Db 1243 AACCGCGCTGGGAAGTCTCAACAGCATCCCGCGCGCTCAGCTCGGTG--AAGCGATG 1299  
 |||||  
 Qy 471 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAsp 490  
 |||||



NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 9707  
LENGTH: 981  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-9707

## Alignment Scores:

Pred. No.: 7, 92e-21 Length: 981  
Score: 295.00 Matches: 91  
Percent Similarity: 43.96% Conservative: 69  
Best Local Similarity: 25.00% Mismatches: 146  
Query Match: 10.33% Indels: 58  
DB: 4 Gaps: 10

US-10-019-783-2 (1-551) x US-09-252-991A-9707 (1-981)

QY 162 ValGlnGluLysGlyProAArgProValLeuProLeuAlaHisGlyAspProSerValPhe 181  
DB 981 CTAGAGGAAGAGGGTCAACCGC---ATCCTCAAGCTGAATATCGCAACCGCGCGCGTTC 925  
QY 182 ProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGly 201  
DB 924 ---GGTTTCGAAGCTCCGAGAAATCTCCAGAGCGTCATCCGCAACCTCCGACCGCC 868  
QY 202 GlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu 221  
DB 867 CAG-----GGCTACAGCGACTCCAAAGGCGCTGTTACGCGCGCGCAAGCAGTATGAG 814  
QY 222 HisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGly 241  
DB 813 TATTACAGCAGACAGCGAGTGGAGCGCTTGGCATCGAGGACATCTACCTCGGCAACGGC 754  
QY 242 GlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsnIle 261  
DB 753 GTGTCGGAACGTGCTGATCTCATCGAGCGCTGTCGAACAC---GGCAGCAGGTG 697  
QY 262 LeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGlu 281  
DB 696 CTGATCCCGCTCCGACTACCGCTGTGGACCGCGCGGTGACCTCGCGCGCGCAAG 637  
QY 282 ValArgHisPheAspLeuLeuProAspLysGlyTyrGluIleAspIleAspSerLeuGlu 301  
DB 636 GCGGTGACTACTGTGGCAGCAGCGCCACTGTGGCGGACCTGGAAGACATCAAG 577  
QY 302 SerIleAlaAspLysAsnThrAlaMetValIleIleAsnProAsnAsnProCysGly 321  
DB 576 GCGAAGATCACCGCGAACACACCGAGCGATGTGATCATCAACCGCAACACCCACCGCG 517  
QY 322 SerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIle 341  
DB 516 GCGGTGATTCCAGGGAAGTCTGGAAGGATGTCGAACTGGCGCGCCGACCAACCTG 457  
QY 342 LeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIlePro 361  
DB 456 GTGCTGTCTCCGACGAGATCTACGACAGATCTCTACACGCGCGCGCTCCAGGCTCC 397  
QY 362 MetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTyr 381  
DB 396 ACCGCTCTCGTGGCGCGGAGCTGCTGCTGCTCACTTCAACGGCGCTGTCCAAATCTAC 337  
QY 382 IleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGln 401  
DB 336 CGGGTGGCGGCTTCCGCTCGGCTGGTGGCGATCTCCGGGCGCCCAAGCAGCGGACAG 277  
QY 401 ----- 401  
DB 276 AGCTATATCAAGGTCTCGACATCTCCCAACATGCGCTGTGGCGCAACGTCCTCCGCGG 217  
QY 402 GluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThr 421  
DB 216 CAGCAGCGGATCCAGACCGCGCTGGCGGCTACGAGCAGCATCAAC----- 172

QY 422 PheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysAla 441  
DB 171 -----GATCTGTCTCTGCGCGC----- 154  
QY 442 IleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLys 461  
DB 153 -----GGCGCTGTGTGAGCAGCAGCAGCGCGCTGGAACTGCTCAACGACATCCCC 100  
QY 462 TyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeu 481  
DB 99 GGCGTCAGTCTGCTG---AAGCGATGGCGCGCTGTACGCTTCCGCGG----- 52  
QY 482 HisLeuLeuGluIleAspAspIleAspPheCysLysLeuAlaLysGluGlu 501  
DB 51 -----ATCGACCCGAAGGTC-----TGCCCGATCCACACGACGAA 16  
QY 502 SerValIleLeu 505  
DB 15 AAGTTCGTCTC 4

## RESULT 13

US-09-107-532A-1579  
Sequence 1579, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynnn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
CORRESPONDENCE ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneske  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 1579:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1173 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...1173  
SEQUENCE DESCRIPTION: SEQ ID NO: 1579:  
US-09-107-532A-1579

## Alignment Scores:

Pred. No.: 1.06e-20 Length: 1173  
 Score: 295.00 Matches: 109  
 Percent Similarity: 43.91% Conservative: 64  
 Best Local Similarity: 27.66% Mismatches: 155  
 Query Match: 10.33% Indels: 66  
 DB: 4 Gaps: 18

US-10-019-783-2 (1-551) x US-09-107-532A-1579 (1-1173)

Qy 170 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 189  
 Db 100 GTGATCAAACTCACCTGGGAGAACCGGAT-----TTTCAACTCCTGAACAT 147  
 Qy 190 AlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnGlySerProAlaGly 209  
 Db 148 GTCAGCAAGCAGGAATAGCAGCGATTAAGAA---GACTTTTCTCACTATACCGGAATG 204  
 Qy 210 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 229  
 Db 205 AGAGGATTAGGAAGTACCGGAAGCGCTGT-----ATATTTCAGCAACAGCGTTAT 258  
 Qy 230 MetLeuSerAlaAsp-----AspValPheLeuThrAlaGlyGlyThrGlnAlaIle 246  
 Db 259 GGATTAACCTTATGATCCGCAACAGGAAGTTTTCAGCACAGTTGTCGCAACGGAAGCGATT 318  
 Qy 247 --GluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArg 265  
 Db 319 GCCTCTCGGTGTATCTGTTTTCGAAGAA-----CGAGCAAGGTCTTATTCACGCG 372  
 Qy 266 ProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPhe 285  
 Db 373 CCTGCTTATTCAGGATATCAGCCA-----TTAGTGGAACTAGCTGTGTCT 417  
 Qy 286 AspLeuIlePro-----AspLysGlyTyrGluIleAspIleAspSerLeuGlu 301  
 Db 418 GAATTGATCCCGATAGACTTCTGACACAGGTTTCTGTCAGCGGACCAATTGAA 477  
 Qy 302 SerIleAlaAspLysAsnThrThrAlaMet-----ValIleIleAsn---ProAsnAsn 318  
 Db 478 CGAGCATTTGAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537  
 Qy 319 ProCysGlySerValTyrSerTyrAspHisSerLysValAlaGluValAlaLysArg 338  
 Db 538 CCTACTGGGACGAGATTCTGCAAAACAAATGAAGCAATCGCTGAGGTTTAAAGAAA 597  
 Qy 339 LeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuVal----- 353  
 Db 598 TATCCGGTATTGTTATAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657  
 Qy 354 LeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSer 373  
 Db 658 ATGTCAPAGTACTATCTTCCA-----GAACAAACAACTCGT 696  
 Qy 374 IleGlySerLeuSerLysSerTyrIleValProGlyTyrArgLeuGlyTyrValAlaVal 393  
 Db 697 ATAACGGGTTATCCAGTACATCCGATGATGATGATGATGATGATGATGATGATGATGAT 750  
 Qy 394 TyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeu 413  
 Db 751 -----TTTGCTCAAAAAGCCCTTATTGACGAACATAATTAAGTCCATCAA 795  
 Qy 414 AsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeu----- 431  
 Db 796 TATCTGTTGATGCGGTACACCAATGTCGCAAAAGCTCGGCTGAGCGCTAATAAAT 855  
 Qy 432 -----GluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeu 446  
 Db 856 GGGGTGAAGATAGCGAGATATGAAGAGCAATATCAAAACACGAGAGACTATCTGCTC 915  
 Qy 447 LysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysPro 466  
 Db 916 CAGCAATTGGCGCCTTATAGGATTGAA-----GTGACC----- 948

Qy 467 HisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGlu 486  
 Db 949 ---CAGCCAAACGGTGCATTTTACTTTGTTGCAAACTCCAGCTGCATTCAA--- 999  
 Qy 487 IleAspAspAspIleAspPheCysCysLysLeuAlaLysGluGluSerValIleLeuCys 506  
 Db 1000 ---ACTGACTCTTGCCAGTTTGTGTAAATTAGCGAACACAGGCAAAAGTTGCTGTATC 1056  
 Qy 507 ProGlySerValLeuGlyMet-----AlaAsnTrpValArgIleThrPheAlaCysVal 524  
 Db 1057 CAGGAATTGCATTTGGCCCAAGAGGAGGTTATATCCGATTTCTTATGTAGTAGCAGC 1116  
 Qy 525 ProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538  
 Db 1117 ATGGAATAATTACATGAGCGCTGCAACCGGATCAAGACTTT 1158

## RESULT 14

US-09-724-623-4

; Sequence 4, Application US/09724623

; Patent No. 6476209

; GENERAL INFORMATION:

; APPLICANT: Glenn, Matthew

; APPLICANT: Lubbers, Mark W

; APPLICANT: Dekker, James

; TITLE OF INVENTION: Polynucleotides, materials incorporating

; TITLE OF INVENTION: them, and methods for using them.

; FILE REFERENCE: 1048U1

; CURRENT APPLICATION NUMBER: US/09/724,623

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 124

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 1107

; TYPE: DNA

; ORGANISM: Lactobacillus rhamnosus

US-09-724-623-4

## Alignment Scores:

Pred. No.: 3.98e-20 Length: 1107  
 Score: 289.00 Matches: 93  
 Percent Similarity: 43.58% Conservative: 63  
 Best Local Similarity: 25.98% Mismatches: 164  
 Query Match: 10.12% Indels: 38  
 DB: 4 Gaps: 13

US-10-019-783-2 (1-551) x US-09-724-623-4 (1-1107)

Qy 170 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 189  
 Db 94 GTCATTGATTGTCTATTGGCCCAACAGAT-----TTTCAACCCCTAAGCGC 141  
 Qy 190 AlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnGlyTyrProAlaGly 209  
 Db 142 ATGTAGCAGCAGCTATTTCGCGCGATTTCAGCTGTGTAATGCCAGTTTCTATACGCAGCA 201  
 Qy 210 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeu-----SerGlnGlyVal 227  
 Db 202 ACCGTTATTCGGATTAAAGCAGCGGATAGTAGCCGATATTTGCCAAGACGGTATT 261  
 Qy 228 ProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGlu 247  
 Db 262 CGTTAT-----GATCATCTGTCATAATCTGTCAACACCGCGCTAAGTTGCTTTGTAT 315  
 Qy 248 ValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGly 267  
 Db 316 GCCTATTTCAGTTTCTTA---NACCCAGCGGATGAGTGTCTGATTCCTGTTCCATAC 372  
 Qy 268 TyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeu 287  
 Db 373 TGGGTTTCTTACGAGAACACAGATTAATTTG---GCGAGCGCGGTGCCACATCTG---GTC 426  
 Qy 288 IleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsn 307





QY 368 IleThrProVal-----LeuSerIleGlySerLeuSerLysSerTrp 381  
Db 319875 ATTCGCGCACTTGCCTCGATTATTAAACCGTTACATTGAATGGTTATCAAAAGCTTAT 319934  
QY 382 IleValProGlyTrpArgLeuGlyTrpValAlaValTyzAspProArg----- 397  
Db 319935 CGAGTTGC-AGCTTCCCGCAAGGCTGGATGATTTTGACGCCCAACATATATGCGAAA 319993  
QY 398 -----LysIleLeuGln 401  
Db 319994 GGTATATTAAGGCTTGATATGCTGCATCAATGCGTTTATGTGCCAACGTACCAATG 320053  
QY 402 GluThrLysIleSerThrSerIleThrAsnTyzLeuAsnValSerThrAspProAlaThr 421  
Db 320054 CAACATGCAATTCAAACCGCACTTGGTGGCTATCAAAAGTATTAT----- 320098  
QY 422 PheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPheLysAla 441  
Db 320099 -----GAATTTATTTACCAGGC----- 320116  
QY 442 IleIleGlyLeuLeuLysGluSerSerGluIleCysTyzLysGlnIleLysGluAsnLys 461  
Db 320117 -----GGTCGATTACTTGACCAACGAAACAAAGCCATGATCTCATCTCAAAATTCCTCA 320170  
QY 462 TyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeu 481  
Db 320171 GGCATTACTTCGGTG---AAACCAATGGGGCGATGATATGTTCCGAAAAATTGATGTG 320227  
QY 482 HisLeuLeuGluGluIleAspAspIleAspPheCysCysLysLeuAlaLysGluGlu 501  
Db 320228 AAAAAAATTC---AATATTCAAGTATGAAATAAATGGTGTGATTTTACTCCGCCAAGAA 320284  
QY 502 SerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrp----- 516  
Db 320285 AAAGTACTACTTGTGCACCGTAAGGATT-----AATTGGCATTACCAGATCAC 320335  
QY 517 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536  
Db 320336 TTCGTATTGTACTCTTCCTTATGTAATCAGCTTGAGAGGCCATTACAAAAATTAGCA 320395  
QY 537 SerPhe 538  
Db 320396 AGATT 320401

Search completed: November 9, 2003, 07:17:20  
Job time : 13168.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 17:12:45 ; Search time 423.595 Seconds  
(without alignments)  
3511.351 Million cell updates/sec

Title: US-10-019-783-2

Perfect score: 2856

Sequence: 1 MATVQSGDVAANGLAFAAA.....LGRKSFQCRKKNSSDDC 551

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2532756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10019783/runat\_07112003\_161642\_19826/app\_query.fasta\_1.1358  
-DB=N Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10019783 @CGN 1 1 458 @runat\_07112003\_161642\_19826 -NCPUS=6 -ICPU=3  
-NO WMAP -LARGEOUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_19Jun03.:

1: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2856	100.0	1910	19	AAV48148
2	2484	87.0	10966	22	AAP32380
3	2142.5	75.0	1660	19	AAV48147
4	1118.5	39.2	1389	24	AB214179
5	998	34.9	1338	24	AB214064
6	998	34.9	1533	21	AA45930
7	992	34.7	1535	21	AA45932
8	872.5	30.5	1481	21	AA45950
9	866.5	30.3	1484	21	AA45953
10	787.5	27.6	1101	21	AA46309
11	787.5	27.6	1182	21	AA45960
12	783.5	27.4	1184	21	AA46314
13	778.5	27.3	1584	21	AA45903
14	663	23.2	858	24	ABN98840
15	656.5	23.0	1377	25	ABO76730
16	642.5	22.5	1844	23	ABL02639
17	637.5	22.3	2754	24	ABN97325
18	637.5	22.3	2782	24	AAH57503
19	637.5	22.3	4412	22	AAH57372
20	567	19.9	5492	23	ABL02638
21	482.5	16.9	1024	25	AB283827
22	360.5	12.6	4403765	22	AA199683
23	360.5	12.6	4411529	22	AA199682
24	354	12.4	48275	21	AA481501
25	354	12.4	34980	21	AA21610
26	354	12.4	34980	21	AA21611
27	354	12.4	837096	21	AA481489
28	351	12.3	349980	22	AA411224
29	350	12.3	1212	25	AB240748
30	348	12.2	1185	18	AA778781
31	348	12.2	1185	25	ABX77257
32	341.5	12.0	1434	22	AA771813
33	341.5	12.0	1434	22	AA772000
34	341.5	12.0	309400	22	AAH68534
35	341.5	12.0	349980	22	AAH68533
36	339	11.9	370	25	ABX20876
37	339	11.9	1257	25	ABX68099
38	336.5	11.8	818	22	AAH20496
39	331	11.6	2365589	24	ABA90521
40	330.5	11.6	48328	22	AAF28540
41	330.5	11.6	2155561	24	ABN71527
42	321.5	11.3	2944528	24	ABA03041
43	318	11.1	349980	24	ABO81845
44	316.5	11.1	1128	24	ABN70754
45	314.5	11.0	1185	25	ABX07293

## ALIGNMENTS

### RESULT 1

AAV48148  
ID AAV48148 standard; cDNA to mRNA; 1910 BP.

AC AAV48148;

DT 27-OCT-1998 (first entry)

DE Nicotianamine aminotransferase 58148.62 molecular weight protein, gene.

KW ds; nicotianamine aminotransferase; plant; iron absorption;  
iron deficiency chlorosis.

OS Gramineae sp.

XX Key Location/Qualifiers  
FH CDS 76..1731

FT /\*tag= a  
 TT /product= "Nicotianamine transferase"

PN EP860499-A2.  
 XX  
 PD 26-AUG-1998.  
 XX  
 XX 19-FEB-1998; 98EP-0102891.  
 PF  
 XX 21-FEB-1997; 97JP-0037499.  
 PR  
 XX (SUMO) SUMITOMO CHEM CO LTD.  
 PA  
 XX Mori S, Nakanishi H, Takahashi M;  
 PI WPI; 1998-439341/38.  
 XX P-PSDB; AAW61643.  
 DR  
 XX New nicotianamine aminotransferase protein and DNA - useful for  
 PT enhancing iron absorption of plant cells  
 PT  
 XX Claim 4; Page 14-15; 17pp; English.  
 PS  
 XX The nicotianamine aminotransferase can be used in a plasmid to transform  
 CC plant cells to produce cells with enhanced iron absorption, and it is  
 CC implied (though not stated) that plants with improved resistance to iron  
 CC deficiency chlorosis in calcareous soils can be regenerated from the  
 CC transformed cells. The gene fragment can be used to detect, amplify  
 CC and/or isolate nicotianamine aminotransferase genes.  
 CC  
 XX Sequence 1910 BP; 462 A; 534 C; 546 G; 368 T; 0 other;  
 SQ

## Alignment Scores:

Pred. No.: 1,71e-219 Length: 1910  
 Score: 2856.00 Matches: 551  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 19 Gaps: 0

US-10-019-783-2 (1-551) x AAV48148 (1-1910)

QY 1 MetAlaThrValArgInSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20  
 DB 76 ATGGCCACCGTACGCCAGAGCAGCGAGTCCGCCGGAACGCCCTTGGCGTGGCGCAGCC 135  
 QY 21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaValAlaAsnGlyLysSerAsnGly 40  
 DB 136 GCGAACGCGCAGAGCAACGCCATGGCGTGGCTGCGCGTGAACGCGCAGAGCAGCGC 195  
 QY 41 HisGlyValAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60  
 DB 196 CATGGCGTGGATGCCGACGCGCAACGCCAAGCAACGCCCATGGCTGGCTGCCGACGCG 255  
 QY 61 AsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAsnGlyHisGlyGluAlaThrAla 80  
 DB 256 AACGCGCAAGACGACGCGCATGCCGAGGCCACTCGCAACGCCACCGCGAGGCCACTGGC 315  
 QY 81 AsnGlyLysThrAsnGlyHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsn 100  
 DB 316 AACGCGCAAGACCAACGCCACCGCGAGAGCAACGCCCATGCTGAGGCGCGCGCAGCGAAC 375  
 QY 101 GlyGluSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAla 120  
 DB 376 GCGCGAGACCAACGACGACATGCCGAGGACTCCGCCGCGCAACGCCGAGACCAACGCCGATGCG 435  
 QY 121 AlaAlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspGly 140  
 DB 436 GCGGCGGCGGCGAGAGAGGAGCGGCTGGAGTGGATTTTCGGGGTGGCCAGGACGCGC 495  
 QY 141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160  
 DB 496 GTGCTGGCGGCGACGGGGCGGAACATGAGCATCCCGCGGATACGGGTACAAGATCAGCGCG 555

QY 161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180  
 DB 556 AGCGTGCAGGAGAGGGCGCGCGCGCTGTGCGCTGGCCACGGGAGCCCGTCCGTG 615  
 QY 181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThr 200  
 DB 616 TTCCCGGCGCTTCCGACCGCGCTCGAGCCCGAGGACGCGCTCGCGCGCGCGTGGCACC 675  
 QY 201 GlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAla 220  
 DB 676 GGCCAGTTCACTGCTACCGCGCGCGCTCCCGCGCGCAGCAGGCGCGTGGCA 735  
 QY 221 GluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAla 240  
 DB 736 GAGCACCTGTGCGAGGCGTGGCGTACATGCTATCGCGCGAGCGCTTCTTCTACCGCC 795  
 QY 241 GlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAlaAsn 260  
 DB 796 GCGGGGACCCAGCGATCGAGTCAATCCCGTGTGGCCACGACCGCGCGCGCAAC 855  
 QY 261 IleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeu 280  
 DB 856 ATTCTGCTCCCGAGCGCGAGGTACCCAACTACGAGCGCGCGCGCTTCAACAGGCTG 915  
 QY 281 GluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeu 300  
 DB 916 GAGTCCGGCATTTGACCTCATCCCGACAAGGGTGGAGATCGACATCGACTCGCTG 975  
 QY 301 GluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCys 320  
 DB 976 GAATCCATCGCGCAAGAACAACACCCCGCATGTGTCATATAAACCACCAACCGCGTGC 1035  
 QY 321 GlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGly 340  
 DB 1036 GGCAGCGTTTACTCTTACGCCATCTGTCAGAGTTCGCGAGGTGGCGAAAGGCTCGGA 1095  
 QY 341 IleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIle 360  
 DB 1096 ATATTGGTGATTGCTGACGAGGTATACGCAAGCTGTTCTGGCGAGCGCGCTTCATC 1155  
 QY 361 ProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSer 380  
 DB 1156 CCAATGGGAGTGTGGGCACATCACCCCTGCTGCTCATTAGGTCTCTGTCCAGTCA 1215  
 QY 381 TrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeu 400  
 DB 1216 TGGATAGTCCCTGGATGGCGCTTGGATGGGTAGCGGTGACGACCCCGAAGATCTTA 1275  
 QY 401 GlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAla 420  
 DB 1276 CAGGAACCTTAAGATCTTACATCAATTACGAATACCTCAATGTCTCGACAGACCAGCA 1335  
 QY 421 ThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLys 440  
 DB 1336 ACCTTCATTTCAGCGAGCTTCTCTCAGATTCTTGAGAACACAAAGAAATTTCTTTAG 1395  
 QY 441 AlaIleIleGlyLeuLysGluSerGluIleCysTyrLysGlnIleLysGluAsn 460  
 DB 1396 GCGATTATTGGTCTGTAAGGAATCATCAGAGATATGCTACAAACAAATTAAGGAAAC 1455  
 QY 461 LysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsn 480  
 DB 1456 AAATACATTACATGCTTCCACAGCCAGAGATCAATGTTGTCTATGGTGAACCTGAAC 1515  
 QY 481 LeuHisLeuLeuGluGluIleAspAspIleAspPheCysLysLeuAlaLysGlu 500  
 DB 1516 TTACATCTTTGGAGAAATAGACATGACATTGATTTTGTGCTGCAAGCTCGCAAAAGAA 1575  
 QY 501 GluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIleThr 520  
 DB 1576 GAATCAGTAATCTTTATGCCCGGAGTGTCTTTGGAAATGGCAAACTGGGTCCGCATTACT 1635

QY 521 PheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleYsSerPheCysGln 540  
 Db 1636 TTTCCTGTGTCCATCTCTCTCAAGATGCTCTCGGAAGATCAATCATTTCTGTCAA 1695  
 QY 541 ArgAsnLysLeuArgAsnSerSerAspAspCys 551  
 Db 1696 AGGAACAAGAGAAGAAATTCGAGCGATTC 1728

RESULT 2  
 AAF32380  
 ID AAF32380 standard; cDNA; 10966 BP.

XX AAF32380;  
 AC  
 XX  
 DT 18-APR-2001 (first entry)

DE Hordeum vulgare L. var. Igri NAAT encoding cDNA SEQ ID NO:1.

XX Hordeum vulgare L. var. Igri; nicotianamine aminotransferase; NAAT;  
 KW NAAT-A; NAAT-B; iron deficiency; gramineous plant; barley; rice;  
 KW mugineic acid biosynthetic pathway; calcareous alkaline soil; ss.

XX Hordeum vulgare.

OS

XX WO200101762-A1.

XX 11-JAN-2001.

PF 04-JUL-2000; 2000WO-JP04425.

XX P-PSDB; AAB69048. AAB69049.

PR 05-JUL-1999; 99JP-0190318.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Mori S, Nakanishi H, Takahashi M, Nishizawa N;

XX WPI; 2001-138030/14.

DR P-PSDB; AAB69048. AAB69049.

XX Gramineous plant, e.g. rice, with tolerance to iron deficiency for

PT growth in calcareous alkaline soil is constructed by transformation

PT with a gene of encoding an enzyme of the mugineic acids biosynthetic

PT pathway

XX Claim 6; Fig 10; 61pp; Japanese.

XX The present invention describes a method for constructing a rice plant  
 CC with improved iron absorability and a tolerance to iron deficiency  
 CC which comprises transferring a gene encoding an enzyme in the mugineic  
 CC acid biosynthetic pathway into a rice plant. The method is for  
 CC constructing gramineous plant e.g. rice with tolerance to iron  
 CC deficiency, which is useful in agriculture in producing new breeds of  
 CC rice plants capable of vigorous growth in calcareous alkaline soil for  
 CC improving crop production. The constructed plant has tolerance to iron  
 CC deficiency, and is therefore capable of vigorous growth in calcareous  
 CC alkaline soil. The present sequence encodes two nicotianamine  
 CC aminotransferases (NAAT), designated NAAT-A and NAAT-B, isolated from  
 CC Hordeum vulgare L. var. Igri (barley), for use in the method of the  
 CC invention.

SQ Sequence 10966 BP; 3082 A; 2422 C; 2366 G; 3096 T; 0 other;

Alignment Scores:

Pred. No.: 1,1e-188 Length: 10966  
 Score: 2484.00 Matches: 550  
 Percent Similarity: 51.30% Conservative: 1  
 Best Local Similarity: 51.21% Mismatches: 0  
 Query Match: 86.97% Indels: 523  
 DB: 22 Gaps: 6

US-10-019-783-2 (1-551) x AAF32380 (1-10966)

QY 1 MetaAlaThrValArgGlnSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20

Db 650 ATGCCACCGTACGCCAGACGCGAGTCCGCGAAGCGCTTGCCTGCGCGAGCC 709  
 QY 21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaValAsnGlyLysSerAsnGly 40  
 Db 710 GCGAAGCGCAAGAGCAACGCCCATGCGTGGCTGCGCGCTGCAACGCGCAAGCAACGCG 769  
 QY 41 HisGlyValAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60  
 Db 770 CATGGCGTGCATGCCGACGCGAAGCAGAGCAACGCCCATGCGTGGCTGCGCGAGCG 829  
 QY 61 AsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAsnGlyHisGlyGluAlaThrAla 80  
 Db 830 AACGGCAAGAGCAACGCCCATGCGGAGGCCACTCGCAACGCCGACGCGGAGGCACTGCG 889  
 QY 81 AsnGlyLysThrAsnGlyHisArgGluSerAsnGlyHisAlaGluAlaAspAlaAsn 100  
 Db 890 AAGCGCAAGAGCAACGCCCATGCGGAGGCCACTCGGAGGCCGCGCGAGCGCGGAGAAC 949  
 QY 101 GlyLysSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAla 120  
 Db 950 GGGGAGAGCAACGAGCATGCCGAGGACTCGCGGGCAACGCCGAGAGCAACGCCGATGCG 1009  
 QY 121 AlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLysAspGly 140  
 Db 1010 GCGCGCGCGCAGAGGAGGCGGTGGAGTGGAAATTCGCGGGTGCCTAAGGACGCG 1069  
 QY 141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160  
 Db 1070 GTGCTGCGCGCAGCGGGCGGAACATGAGCATCCGCGCGATACGTTACAGATCAGCGCG 1129  
 QY 161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180  
 Db 1130 AGCGTGCAGAGAGGCGCGCGCGTGTGTCGCGTGTGTCGCGCGCGGAGCGCGTCCGTG 1189  
 QY 181 PheProAlaPheArgThrAlaValAlaGluAspAlaValAlaAlaValAlaArgThr 200  
 Db 1190 TTCCCGCGCTTCGCGACGCGCGTTCGAGGCGGAGAGCGCGTGCCTCCCGCGCGTCCGAC 1249  
 QY 201 GlyGluPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArg----- 216  
 Db 1250 GGCCAGTTCAACTGCTACCGCGCGCGTCCGCGTCCCGCGCGCGAGGTAACAACAA 1309  
 QY 216 ----- 216  
 Db 1310 CAACAACAAGAAACAATTCCTTTTCGCGTGTGTCGCGCGCAATCCATGATGCG 1369  
 QY 216 ----- 216  
 Db 1370 CATGTGCGGCTTTCACGTGTGCGTCCGTCACCGTTCCTCTCTCTCCCTACGCGC 1429  
 QY 217 -----SerAlaVal 219  
 Db 1430 ATGAGAAATCTGACCTTCTCCACCTTATACCAAAACAAACAAACACACGCGCGTG 1489  
 QY 220 AlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThr 239  
 Db 1490 GCAGAGCACTGTGCGAGGCGGTGCGGTACATGCTATCGCGCGACGACGCTCTCTCACC 1549  
 QY 240 AlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAla 259  
 Db 1550 GCGCGCGGACCCAGCGCATCGAGGTCAATAATCCCGGTGCTGCGCGCAGACCGCGCGCG 1609  
 QY 260 AsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArg 279  
 Db 1610 AACATTTCTCTCCCGAGCGAGGTACCCAACTACGAGCGCGCGCGGTTCACAGG 1669  
 QY 280 LeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSer 299  
 Db 1670 CTGGAGGTCCGCGCATTTCCACCTCATCCCGCAACAGGGGTGGAGATCGACATCGACTCG 1729  
 QY 300 LeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnPro 319

Db 1730 CTGGAATCCATCGCGCAAGAACACACACCGCCATGTGTCATCAATAACCCCAACACCCG 1789  
Qy 320 CysGlySerValTyrSerTyrAspHisLeuSerLys----- 331  
Db 1790 TGGCGCAGCGTTTACTCTCAGACCACTGTCCAGGTTTCACATCCTTTGGCCTTGCTGA 1849  
Qy 331 ----- 331  
Db 1850 ATATGGAATTCAGTTCACTGACGACCTGCTGAATCTTTTGGCAATCGCATACTGACTGATG 1909  
Qy 332 -----ValAlaGluValAlaAlaGlySerLeuGlyIleLeuValIleAlaAspGlu 347  
Db 1910 TTGCTCAATTAGTTCGCGAGGTGCGAAAGGTCGGAATTTGGTGTGTTGCTGACGAG 1969  
Qy 348 ValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHis 367  
Db 1970 GTATACGGCAAGCTGGTTCTTGGCGAGCGCCCGTTTCATCCCAATGGGAGTGTTTGGGCAC 2029  
Qy 368 IleThrProValLeuSerIleGlySerLeuSerLysSerThrIleValProGlyTyrArg 387  
Db 2030 ATCACTGCTGCTGTCATAGGTCCTGTGCAAGTCATGATGATGCTGCTGATGCGCG 2089  
Qy 388 LeuGlyTyrValAlaValTyrAspProArgLysIleLeuGlnGluThrLys----- 404  
Db 2090 CTTGGATGGTAGCGGTGTCAGACCCAGAAAGATCTTACAGGAACTAAAGTACTTTAAA 2149  
Qy 404 ----- 404  
Db 2150 TCTCTATATCATCTTTTCAAATGCTACTAAGGTGATTAATTAGTACTACTGTACAAATAT 2209  
Qy 405 -----IleSerThrSerIleThrAsnTyrLe 413  
Db 2210 ATTGCTAAATTTGCTACTGACATTTTGTGGTAGATCTCTACATCAATTACGAATTACCT 2269  
Qy 413 uAsnValSerThrAspProAlaThrPheIle----- 423  
Db 2270 CAATGCTCGACAGACCCAGCAACCTTCATTTCAGGTCAGTCTTTGGTATTACCTCGTTT 2329  
Qy 423 ----- 423  
Db 2330 CAAGAAATAAGTCTTTGGTATTACTCTCTGCTGCTATTTTGTCCGGTCCCTATGT 2389  
Qy 423 ----- 423  
Db 2390 TGTAGGCAGCCACGTGCATGTCAAGTGACCGTTTTTTCACATTAAGTTTGAAGTCAAA 2449  
Qy 423 ----- 423  
Db 2450 GTCAGACACATACACTGTAGTTATTTTACCTTTGTTGTTTGTTCATCCGATAAAATAAAA 2509  
Qy 423 ----- 423  
Db 2510 AAATACAAAACCTGAACCTACTGTGTAATATAACCACTGTTCTTACAAGATATACATGAT 2569  
Qy 423 ----- 423  
Db 2570 TGCATATGGCGATGCCATATTTCTTTGGGTCAAGTATGACATGTTGGAAACCTTTT 2629  
Qy 423 ----- 423  
Db 2630 AGAATATAGATACATGTACTGATATACCAATTTATTAAGAAATTCATATTTTGATA 2689  
Qy 423 ----- 423  
Db 2690 TCCTTGATGGTATTCTCTTGTGATTACACAGATTTTACTTGTGGTTTTTGTACTATC 2749  
Qy 424 -----GlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLys 441  
Db 2750 AAATGTTTCAGCAGCTCTTCTCAGATCTTTGAGACACAAAGAGATTTCTTTAAGG 2809  
Qy 441 laileileGlyLeuLeuLysGluSerGluIleCysTyrLysGlnIleLysGluAsnL 461  
Db 2810 CGATTATTGGTCTGCTAAAGGAATCATCAGATATGCTACAAACAAATAAAGGAAAAACA 2869

Qy 461 yeTyrIleThrCysProHisLysProGluGlySerMetPheValMet----- 476  
Db 2870 AATACATTACATGTCTCACAAGCCAGAAGGATCAATGTTTGTCTATGTAAGCCTATTTT 2929  
Qy 476 ----- 476  
Db 2930 GTGAAGTAAAAAATCTTAGGAGTGTGAGTAATCATAACTTATTTATATAGGATTAAT 2989  
Qy 476 ----- 476  
Db 2990 CTGGGACCGAATGATCCAAACATAATTACTTCAAATTTCAAATTTACAAATTTCTTCC 3049  
Qy 476 ----- 476  
Db 3050 GTACATATTTTGAAGATGCATGATTTTAAGAATAATGACGAGAGCTAAAGTTATGCTA 3109  
Qy 476 ----- 476  
Db 3110 CGACTAATCATCTGGATATCCTTTTGTCCATCTTTTGTATACTGTGGAATGTTAATGTT 3169  
Qy 476 ----- 476  
Db 3170 CAAATCATATTACACAAATATCCATGCTAGTTTCTAGAAAGATGATTTTCTGTAA 3229  
Qy 477 -----ValLysLeuAsnLeuHisLeuGlu 485  
Db 3230 CCATGAACCTCGTATTAACTTCCATGTAACAGGTGAACTGAACTTACATCTTTGGAG 3289  
Qy 486 GluIleAspAspAspIleAspPheCysCysLysLeuAlaLysGluLysSerValIleLeu 505  
Db 3290 GAAATAGACATGACATTTGCTGCAAGCTCGCAAAAGAGAAATCAGTAATCTTA 3349  
Qy 506 CysPro----- 507  
Db 3350 TGCCCGAGTAGGAATCCATGTTGATTTTGTGCTATATGAAGTCTTATCAATTTCCG 3409  
Qy 507 ----- 507  
Db 3410 AGATGACTATACATATAAATGATTACCATATTTATGGTCAGAAATTTGATACAGTGTAG 3469  
Qy 507 ----- 507  
Db 3470 AATATTCTGTGAAGACTTTTTPAACAAATATCTGTGAAGACTAGATATCATGTACTTC 3529  
Qy 507 ----- 507  
Db 3530 TCCTTGTCTTCTTGACCTGATGCTCTGTCACATGTTGCTGCTCCACAAAAAATAGC 3589  
Qy 507 ----- 507  
Db 3590 AAGCACATGTTTCAAATAATTTGTTAAATAATAATTAGCCCTTTAATTTATATGTTCTA 3649  
Qy 507 ----- 507  
Db 3650 TTTTGAGATATTTTGTAGTCCAACTTATATATTTGTGACTATTTCTCAAAACAAACTT 3709  
Qy 508 -----GlySerValLeuGlyMetAlaAsnTrpValArgIle 519  
Db 3710 ATATATGTGCTCTCTCAAATGTAGGAGTGTCTTGGAAATGGCAAACTGGGTCCGCATT 3769  
Qy 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyValGlyIleLysSerPheCys 539  
Db 3770 ACTTTGCTTGTGTTCCATCTTCTTTCAAGATGTCCTCGGAAGATCAATCATCTTCTGT 3829  
Qy 540 GlnArgAsnLysLysArgAsnSerSerAspAspCys 551  
Db 3830 CAAAGGAACAAGAAGAGAAATTCGAGCGATGATTGC 3865

RESULT 3

AAV48147

ID AAV48147 standard; cDNA to mRNA; 1660 BP.

XX

```
AC AAV48147;
XX
XX DT 27-OCT-1998 (first entry)
XX DE Nicotianamine aminotransferase 49564.15 molecular weight protein, gene.
XX KW ds; nicotianamine aminotransferase; plant; iron absorption;
XX KW iron deficiency chlorosis.
XX OS Gramineae sp.
XX PH Key Location/Qualifiers
XX FT CDS 62..1447
XX FT /tag= a
XX FT /product= "Nicotianamine aminotransferase"
XX PN EP860499-A2.
XX PD 26-AUG-1998.
XX PF 19-FEB-1998; 98EP-0102891.
XX PR 21-FEB-1997; 97JP-0037499.
XX PA (SUMO ) SUMITOMO CHEM CO LTD.
XX PI Mori S, Nakanishi H, Takahashi M;
XX WPI; 1998-439341/38.
XX DR P-PSDB; AAW61642.
XX CC New nicotianamine aminotransferase protein and DNA - useful for
XX PT enhancing iron absorption of plant cells
XX PS Claim 4; Page 12-13; 17pp; English.
XX CC The nicotianamine aminotransferase can be used in a plasmid to transform
CC plant cells to produce cells with enhanced iron absorption, and it is
CC implied [though not stated] that plants with improved resistance to iron
CC deficiency chlorosis in calcareous soils can be regenerated from the
CC transformed cells. The gene fragment can be used to detect, amplify
CC and/or isolate nicotianamine aminotransferase genes.
XX SQ Sequence 1660 BP; 423 A; 442 C; 430 G; 365 T; 0 other;

Alignment Scores:
Pred. No.:      2,34e-162      Length:      1660
Score:          2142.50        Matches:     415
Percent Similarity: .93.5%    Conservative: 20
Best Local Similarity: .89.25% Mismatches:   21
Query Match:      75.02%       Indels:      9
DB:              19           Gaps:        3

US-10-019-783-2 (1-551) x AAV48147 (1-1660)

QY 89 GluSerAsnGlyHisAlaGlu-----AlaAlaSplaaSnglyGluSerAsnGluHis 106
Db 71 CGAGCAACCGGCCACGGCAGCGCCGCCGCCGCCGCCAACGGCAAGACAACGGGCAC 130
QY 107 AlaGluaSPserAlaAalaSnGlyGluSerAsnGlyHisAlaAalAAlaAaLgUglU 126
Db 131 GCC-----GCCCGCCGAACGGCAAGCAACGGGCACC GGCGCGGCGG----- 175
QY 127 GluGluaValglUTrpaSnPheAlaClAlaylaspGlyValLeuAlaAlaThrGly 146
Db 176 -----GCCGTGGATGGTAATTTCGCCCGGGCACGACGCCCATCTTGGCGACGACGGG 229
QY 147 AlaAsnMetSerIleArgAlaAttleArgTrvrltylleSerAlaserValGlnGluLysGly 166
Db 230 GCGAAGAACAGCATCCGGCGGNATCGGTACAAGATCAGCGGAGCGTGAGAGACGCGG 289
QY 167 ProArgProValLeuProLeuAlaHisGlyASPProSerValPheProAlapHeArgThr 186
```

```

QY 547 SerSerAspAspCys 551
DB 1430 TCTATAATGGTTGT 1444

RESULT 4
ID ABZ14179 standard; DNA; 1389 BP.
XX AC ABZ14179;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1984.
XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX OS Arabidopsis thaliana.
XX PN WO200216555-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-US26685.
XX PR 24-AUG-2000; 2000US-227866P.
XX PR 26-JAN-2001; 2001US-264647P.
XX PR 22-JUN-2001; 2001US-300111P.
XX PA (SCRI ) SCRIPPS RES INST.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Harper JF, Krebs J, Wang X, Zhu T;
XX DR WPI; 2002-304127/34.
XX PT Identifying a stress condition to which a plant cell has been exposed
XX PT and producing plants with increased tolerance to these abiotic stresses
XX PS Claim 144; SEQ ID NO 1984; 577pp + Sequence Listing; English.
XX CC The invention relates to identifying a stress condition to which a plant
XX CC cell has been exposed, comprising:
XX CC (a) contacting nucleic acid representative of expressed polynucleotides
XX CC in the plant cell with an array or probes representative of the plant
XX CC cell genome; and
XX CC (b) detecting a profile of expressed polynucleotides in the plant cell
XX CC characteristic of a stress response. The method is useful in the
XX CC production of transgenic plants, cells and seeds and in producing plants
XX CC with increased tolerance to abiotic stress. The present sequence is that
XX CC of an Arabidopsis thaliana stress regulated gene (ABZ121196-ABZ17574) used
XX CC in methods of the invention.
XX CC Note: The sequence data for this patent is not represented in the printed
XX CC specification but is based on sequence information supplied to Derwent by
XX CC the European Patent Office.
XX SQ Sequence 1389 BP; 389 A; 322 C; 337 G; 341 T; 0 other;

Alignment Scores:
Pred. No.: 2,41e-80 Length: 1389
Score: 1118.50 Matches: 207
Percent Similarity: 65.94% Conservative: 97
Best Local Similarity: 44.90% Mismatches: 140
Query Match: 39.16% Indels: 17
DB: 24 Gaps: 3

US-10-019-783-2 (1-551) x ABZ14179 (1-1389)

QY 89 GluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluHisAlaGlu 108
DB 7 GAAGAACCAACACCGCAATCTAGCGGTTCCCGGTTTAAACTGAGAAAGAGCCCAATA 66

```



```

Db 1096 CTGAGTCTTGCACTTACTTATTGACAAAGTTGGAGCTGTCAATTGATGATAAATCAAG 1155
QY 489 AspAspIleAspPheCysLysLeuAlaLysGluSerValIleLeuCysProGly 508
Db 1156 GACGATATAGATTCTTGGTAAACTGCGCAGAGAGAGAGAACTCTGGTTTCTACCAAGG 1215
QY 509 SerValLeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeu 528
Db 1216 GATGCTCTGGTTTGAAGAACTGGATGAGGATAACCATCGGAGTCGAAGCTCATATGCTT 1275
QY 529 GlnAspGlyLeuGlyArgIleLysSerPheCysGlnArgAsnLysLysArgAsnSerSer 548
Db 1276 GAGGATGCACTGAGAGACTGAAGGTTTCTGACAGCTCATGTCACCAAGAAGACAGACACA 1335
QY 549 Asp 549
Db 1336 GAA 1338

```

## RESULT 5

ABZ14064  
ID ABZ14064 standard; DNA; 1338 BP.

AC ABZ14064;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1869.

KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

PN WO200216655-A2.

PD 28-FEB-2002.

PF 24-AUG-2001; 2001WO-US26685.

PR 24-AUG-2000; 2000US-227866P.

PR 26-JAN-2001; 2001US-264647P.

PR 22-JUN-2001; 2001US-300111P.

PA (SCRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;

DR WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed

PS Claim 144; SEQ ID NO 1869; 577pp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant

CC cell has been exposed, comprising:

CC (a) contacting nucleic acid representative of expressed polynucleotides

CC in the plant cell with an array or probes representative of the plant

CC cell genome; and

CC (b) detecting a profile of expressed polynucleotides in the plant cell

CC characteristic of a stress response. The method is useful in the

CC production of transgenic plants, cells and seeds and in producing plants

CC with increased tolerance to abiotic stress. The present sequence is that

CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used

CC in methods of the invention.

CC Note: The sequence data for this patent is not represented in the printed

CC specification but is based on sequence information supplied to Derwent by

CC the European Patent Office.

## Alignment Scores:

```

Pred. No.: 1,06e-70 Length: 1338
Score: 998.00 Matches: 194
Percent Similarity: 65.51% Conservatives: 89
Best Local Similarity: 44.91% Mismatches: 139
Query Match: 34.94% Indels: 10
DB: 24 Gaps: 3
US-10-019-783-2 (1-551) x ABZ14064 (1-1338)
QY 111 AlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaAlaGluGluGluAlaVal 130
Db 4 GCAGCAACGGAGTGTACCAACTGTAAACGCAACGCCAATGTT----- 45
QY 131 GluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMetSer 150
Db 46 ---TGGCGGTCAAAGGA-----AACGGTGCA---ACGAGTCATCGCGCGGGTGACG 93
QY 151 IleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArgProVal 170
Db 94 TTGAGAAAGCTTGGCTTTTGGGATGTTTAAAAAATGCACCATGAACAGTGGAAAGACCAT 153
QY 171 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 190
Db 154 TTGTTCCCACTCCCGGAGAGCCCTCCGCCATTCCAACTTCAGGACTTGCCTGGGAGCC 213
QY 191 GluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyVal 210
Db 214 GAGGAAGCCGTTGCCGAGTCGACGCTCCGGCATGGCTAACTTACGACCCAGCCCT 273
QY 211 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMet 230
Db 274 GGAGTTTTCAGGCTAGAAAGGGCGGTGGCTGAATATTAAACGGAGAACTTCGAGCAAG 333
QY 231 LeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 250
Db 334 CTGAGGCGCGAGGATGTGTATATACCGGAGGATGTAAACCAAGCCATAGAGATCGTGA 393
QY 251 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 270
Db 394 GATTCTCTTCCCGGAAATCCATCCGCCAACATTTACTTCCAAAGGCGCGGATATCTCAC 453
QY 271 TyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProAsp 290
Db 454 TAGCATGCTCGTGTCTCTATAGCGGCTCGAGATTCGCAATACGATTTCTCCCGAG 513
QY 291 LysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrAla 310
Db 514 AGTGATTGGGAAATCAATCTCGATGGCTCGAGGCGGCTCGGATGAGAATACCGTCGCA 573
QY 311 MetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330
Db 574 ATGGTAATCATCAACCCCAACATCATGTGTGAACGCTCTACACCTACGACCATCTCAAC 633
QY 331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350
Db 634 AAGGTCGGGAGATGGCTAGAAACTCGGTATAATGATAATATCCGACCAAGTATATGAT 693
QY 351 LysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIleThrPro 370
Db 694 CATGTTGTATATGGAGACAAAGCCCTTTATTCCTCCATGGGGAAGTTGCATCAATAGTCCG 753
QY 371 ValLeuSerIleGlySerLeuSerLysSerTyrIleValProGlyTrpArgLeuGlyTrp 390
Db 754 GTGATCAGCTCGGATCCATATCCAAAGGATGGGTCAACCCAGGCTGGAGAGTTGGCTGG 813
QY 391 ValAlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSerIleThr 410
Db 814 ATCGCCATGAACGATCCTAATGGTATCTTTGTATCTACAGGGGTAGTTCAAGCAATAGAG 873
QY 411 AsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIle 430
Db 874 GATTTCCTGATTTAACTCCACAGCCCTTCATTTATCTCCAGAGACATTTCTCTGATATA 933

```





```
Db      818 GTGATCAGCTCGATCCATATCAAGGATGGTCAACCCAGGCTGGAGTGGCTGG 877
Qy      391 ValalaValtyrAspProArgLysIleLeuGlnGluThrIleSerIleThr 410
D5      878 ATCCGATGAACGATCCATATGATGTTATCTACAGGGTAGTTCAGCAATAGAG 937
Qy      411 AsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaLeuProGlnIle 430
D5      938 GATTTCTTGAATTAATCACTCCACGCTTCATTTATCTCCAGGAGCATTCTGATATA 997
Qy      431 LeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSer 450
D5      998 TTGAGAAAACACCTAAAGAGTCTTCGAGAAGAAGATCAAGCCATGAGACGCAACGTC 1057
Qy      451 GluLeuCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLysProGlu 470
D5      1058 GAGCTTTCATGTGAGAGCTCAAGGATATCTCTGTCTCTTTGTCCCAAGAACCCGAA 1117
Qy      471 GlySerMetPheValMetValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAsp 490
D5      1118 TCTTGTCTTATTTATGTTGATGCTGACATCATGTTGATATATATCAAAATGAT 1177
Qy      491 IleAspPheCysCysLysLysLeuAlaLysGluSerValIleLeuCysProGlySerVal 510
D5      1178 TTTGATTTCTGCACAAAGTCTAGTGTAGTGAGGAGAGTCTTATCCTTATACCAAGAGTGGCT 1237
Qy      511 LeuGlyMetAlaAsnTyrValArgIleThrPheAlaCysValProSerSerLeuGlnAsp 530
D5      1238 CTAGGGCAGAGATTTGGTGAGGATATCGATAGGACCGACGATCAGTGGTACAGAA 1297
Qy      531 GlyLeuGlyArgIleLysSerPheCysGlnArgAsn 542
D5      1298 ATATTGACAGACTAAAGGTTTCTATGATCGTCAT 1333

RESULT 7
AAC35232
ID AAC35232 standard; DNA; 1535 BP.
AC AAC35232;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9451.
XX
XX Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 24-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131443.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144884.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 21-JUL-1999; 99US-0145086.
```

```
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145195.
PR 23-JUL-1999; 99US-0145197.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145277.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0151930.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154033.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 3,82e-70
Score: 992.00
Percent Similarity: 65.28%
Best Local Similarity: 44.68%
Query Match: 34.73%
DB: 21
Gaps: 3

US-10-019-783-2 (1-551) x AAC35232 (1-1535)

QY 111 AlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaAlaAlaGluGluAlaVal 130
Db 68 GCGAGCAACGGAGTTTACCAACTGTAAACGCAACGCAATGTT----- 109
QY 131 GluTirAsnPhelaAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMetSer 150
Db 110 ---TGGCGGTTCAAGGA-----AACGGTGCA---ACGAGTCATCGCGGCGGTGACG 157
QY 151 IleArgAlaIleArgTyrIlySileSerAlaSerValGlnGluGlyGlyProArgProVal 170
Db 158 TTGAGAAAGCTGCTTTTGGGATGTTTAAAAAGTCACCATGAACAGTGGAAAGACCAAT 217
QY 171 LeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGluAla 190
Db 218 TTGTTCCCACTCCCGGAGCGCTCCGCCCATTCACATTCAGGACTTGGCCGGAAGCC 277
QY 191 GluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGlyVal 210
Db 278 GAGGAAGCGGTGGCGACGCTGCACTCCGGCATGCTTAACCTTACGCCACCCACCCCT 337
QY 211 GlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyrMet 230
Db 338 GGAGTTTTCAGGCTAGAGGGCGGTGGCTGAATATTAAACGGAGAACTTCGACGAAG 397
QY 231 LeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIle 250
Db 398 CTGAAGCGCGAGGATGCTATATATACCGGAGGATGTAAACCAAGCCATAGATCGTGATA 457
QY 251 ProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsn 270
Db 458 GATTCCTTTCGCGAAATCCATCCGCCAACATTTACTTCCAGGCGCGGATATCTCTCAC 517
QY 271 TyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIleProAsp 290
Db 518 TAGCATCTCGTCTGCTCTATAGCGGCTCGAGATTCGCAATACGATTTCTCCCGAG 577
QY 291 LysGlyTirpGluIleAspIleAspSerLeuGluSerIleAlaAspIlyAsnThrThrAla 310
Db 578 AGTGATTGGAAATCAATCTCGATGCCCTCGAGGCGGTGGGATGAGATACCGTCGCA 637
QY 311 MetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeuSer 330
Db 638 ATGGTAATCATCAACCCCAACAATCCATGTGAAACGCTACACCTACCATCTCTCAAC 697
QY 331 LysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluValTyrGly 350
```



[illegible][illegible]





```
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 17-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145119.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154773.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0155653.
PR 28-SEP-1999; 99US-0156458.

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158222.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 4,23e-60
Score: 866.50
Percent Similarity: 61.59%
Best Local Similarity: 41.92%
Query Match: 30.34%
DB: 21
Length: 1484
Matches: 179
Conservative: 84
Mismatches: 145
Indels: 19
Gaps: 6

US-10-019-783-2 (1-551) x AAC39233 (1-1484)
Qy 128 GluAlaValGluTrpAsnPheAlaGly-----AlaLysAspGlyValLeuAlaAla 144
Db 52 AAGTGCATTGATTGGCAATTCAGCGGAGCGGCGCAAGATGCTCCCGCGCTCC 111
Qy 145 ThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGlu 164
Db 112 TTAGGCTCATATACCTCT-----GCACCTATGCACTGCGGATCCT 153
Qy 165 LysGlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPhe 184
Db 154 CATGGC---AAACCCATTGTGCCCGCCACGAAAT-----GAGATCCGTG 192
Qy 185 ArgThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsn 204
Db 193 GAGACCAAGCAATACAGCCGCAAAAGAGCTTGTAAAGCTGTTCTTATGGCTCGGGAAC 252
Qy 205 CysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSer 224
Db 253 GCCTATGCTCTAGCTTAGGCTCGCGCGCCGCAAAAGTCCGCTAGCAGATATCTAAAC 312
Qy 225 GlnGlyValProTyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGln 244
Db 313 CAAGGTCTTCCAAAGAGCTTACAGCAGATGACGTGTTATACCCCTTGATGCGCAACAA 372
Qy 245 AlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuPro 264
Db 373 GCTATTGAGCTCGCGGTAGACATTCGCTAAACCGAAA---GCCAACGTTTGTCTTCG 429
```

```
QY 265 ArgProGlyTyrProAsnTyrGluAlaAargAlaAaPheAsnArgLeuGluValArgHis 284
DB 430 AGTCCCGGCTTCCCATGGACCTAGTCGGTCCATCTACAGAACCTTGAGTCCGCCAC 489
QY 285 PheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSerLeuGluSerIleAla 304
DB 490 TATAATTTCCTCCAGAAAGAACTTGAATTCGAAATCGACTTTGATAGCGTCGCGCGCTCGTG 549
QY 305 AspLysAsnThrThrAlaMetValIleAsnProAsnAsnProCysGlySerValTyr 324
DB 550 GACGAGAACACATTGGCATATTTATATCAACCCCAACCCCAACCCCAATGTAACACCTAC 609
QY 325 SerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIle 344
DB 610 TCCTAGGCTCATCTCAACACGCTCGCTGATTTGGCTAAGGAACCTCAAGATTATGGTGTT 669
QY 345 AlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyVal 364
DB 670 TCTGACGAGGTTTATAGATGACACTCTTTGGGAGTAACCCCTTTGTCCCTATGGGAAA 729
QY 365 PheGlyHisIleThrProValLeuSerIleGlySerLeuSerLysSerTrpIleValPro 384
DB 730 TTCCTCGTCGATCGTACCAAGTGGTTACACTCGATCCATATCAAGGATCGAAGTCCCA 789
QY 385 GlyTrpArgLeuGlyTyrValAlaValTyrAspProArgLysIleLeuGluThrLys 404
DB 790 GGATGGCGAAGTGGTGGCTCACCTACATGATCTAGACGGTGTCTTCAGAAACACCAAG 849
QY 405 IleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGln 424
DB 850 GTCCTACAGTGTCTAAGATTTCTCGATATAACAATAACCTCCGACAGTTATCCAG 909
QY 425 AlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleGly 444
DB 910 CGCGCTATCTCGACATCTTGAGAAACCTCTCAAGAGTCTTCGATAAGAGCGCAGAGT 969
QY 445 LeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThr 464
DB 970 TTTCTGAAAGATAAAGTAGAATTTGGTTATTTAAGCTCGAGTACATCTCGCTCACT 1029
QY 465 CysProHisLysProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuLeu 484
DB 1030 TGCTACATGAACCCGAGCGCTGACCTTCTATGGACCGAGCTTGATTATCGAGCTTT 1089
QY 485 GluGluIleAspAspAspIleAspPheCysLysLeuAlaLysGluGluSerValIle 504
DB 1090 GTGACATCGAAGACGATCAAGACTTTTGCAATAAGCTTGTAAAGAAAGAAACCTCGTC 1149
QY 505 LeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysVal 524
DB 1150 GTTTTACAGGATGTGATTCAGTCAGACAGTGGTTGAGGCATTCATCGATCGGAG 1209
QY 525 ProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGlnArg-----Asn 542
DB 1210 ACTCCGCTATTGGAGTGCATTGGAAGATTGAAGAGCTTCTGCGATCGCCATTCACAC 1269
QY 543 LysLysArgAsnSerAsp 549
DB 1270 AAAAAGCTCCCTCAAGAC 1290
RESULT 10
AAC46309
ID AAC46309 standard; DNA; 1101 BP.
XX
AC AAC46309;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 49668.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW
```

```
KW XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN EP1033405-A2;
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
```

PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0156559.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 20-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 14-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145228.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 23-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 23-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 23-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149920.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
Alignment Scores:			
Pred. No.:	6.22e-54	Length:	1101
Score:	787.50	Matches:	152
Percent Similarity:	86.2%	Conservative:	70
Best Local Similarity:	45.37%	Mismatches:	110
Query Match:	27.57%	Indels:	3
DB:	21	Gaps:	2
US-10-019-783-2 (1-551) x AAC46309 (1-1101)			
QY 217	SerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspVal	236	
Db 82	AGTGGCTAGCAGATATCTAAACCAAGCTCTCCAAAGACCTACCCGACATGACGTG	141	
QY 237	PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThr	256	
Db 142	TTTATGACTCTGGGTGCAACAACTATTGAGCTCGCGGTAGACATTCCTCGTAAACCG	201	
QY 257	AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaLa	276	
Db 202	AAA--GCCACAGCTTTGCTTCAGAGTCCGGCTTCCCATGGGACCTAGTCGCTCCATC	258	
QY 277	PheAsnArgLeuGluValAlaArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp	296	
Db 259	TACAAGACCTTGAGTCCGCCACATATATTTCTTCAGAAAGAACTTTGAATATCGAC	318	
QY 297	IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro	316	

Db 319 TTTGATAGCGTCGAGCGCTCGTGACGAGACACACATTTGCCATATTTATATCAACCCC 378  
Qy 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336  
Db 379 CACAACCCCAATGTAACACTACTCGAGCGCTCATCTCAACAGCTGGCTGAACCTGGCT 438  
Qy 337 LysArgLeuGlyLeuValAlaAspGluValTyrGlyLysLeuValLeuGlySer 356  
Db 439 AAGAACTCAAGATATGTTGGTTCTTGACGAGGTTTATAGATGGACACTCTTTGTAGT 498  
Qy 357 AlaProPheLeuProMetGlyValPheGlyHisLeuThrProValLeuSerLysGlySer 376  
Db 499 AACCCCTTTGTTCTATGGGAAATCTCGTCGATCGTACAGTGGTTACACTCGCATCC 558  
Qy 377 LeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspPro 396  
Db 559 ATATCAAGGGATGAAAGTCCAGGATGGCACTGGTTGGCTCAGCTACAGTACATGACTA 618  
Qy 397 ArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 416  
Db 619 GACGGTGCTTCAGAAACACCAAGCTCTTCAAGCTGCTCAAGATTTTCTCCAGATAAAC 678  
Qy 417 ThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGlu 436  
Db 679 AATAACCCCTCGCAGATTCACAGCGCTATTCCTGACATCTTGGAGAAAACCTCCTCAA 738  
Qy 437 AspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGln 456  
Db 739 GAGTTTITGATAGACGACAGAGTTTCTGAAGATTAAGTAGAATTTGTTTATTCTAAG 798  
Qy 457 IleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet 476  
Db 799 CTCAGTACATTCCTAGCTCCTTGTATGATGAACCCGAGCTGACCTTCTTATGG 858  
Qy 477 ValLysLeuAsnLeuHisLeuGluGluIleAspAspAspIleAspPheCysCysLys 496  
Db 859 ACCGAGTTGATTTATCGAGTTTGTGACATCGAAGACGATCAAGACTTTTGCATAG 918  
Qy 497 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrp 516  
Db 919 CTTGCTAAAGAAAGAAACCTCGCTTTTACCAGGATTCATTCAGTACAGAAAGACTGG 978  
Qy 517 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536  
Db 979 TTGAGGATTCATCGATATGGAGACTCCGGTATGGAGATGCATTCGGAAGATTGAAG 1038  
Qy 537 SerPheCysGlnArg-----AsnLysLysArgAsnSerSerAsp 549  
Db 1039 AGCTTCTCGATCGCATTCCTCAACAAAAAGCTCCCTCAAGAC 1083  
RESULT 11  
AAC45560  
ID AAC45560 standard; DNA; 1182 BP.  
XX AAC45560;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 46941.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
FN EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.

```

PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 02-AUG-1999; 99US-0145389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148344.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 6,82e-54
Score: 787.50
Percent Similarity: 66.27%
Best Local Similarity: 45.37%
Query Match: 27.57%
DB: 21

US-10-019-783-2 (1-551) x AAC45560 (1-1182)

QY 217 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspVal 236
Db 11 AGTGCCTAGCAGAGTATCTAACCAAGGTCTTCAAGAGCTTACCCAGATGACGTG 70
QY 237 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThr 256
Db 71 TTTATGACTCTGGATGCAACACAGCTATTGAGCTCGCGGTAGACATTCTCGCTAAACCG 130
QY 257 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 276
Db 131 AAA--GCCAAGCTTTTGCTTCGAGTCCCGGCTTCCCATGGGACCTAGTCGCTCCATC 187
QY 277 PheAsnArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp 296
Db 188 TACAAGACCTTTGAGGTCCGCCACTATATTTCTTCGAGAAAGAACTTTGAAATCGAC 247
QY 297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 316
Db 248 TTTGATAGCGTCGAGCGCTCGTGACGAGACACATTTGCCATATTTATATCAACCCC 307
QY 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336
Db 308 CACACCCCAATGGTAACACTACTCCGAGGCTCATCTCAACACGCTGGTGAATGGCT 367
QY 337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
Db 368 AAGGAACCTCAAGATTATGTTGTTCTGACGAGGTTTITTAGATGACACACTCTTTGGTAGT 427
QY 357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySer 376

```

Db 428 AACCTTTTGTCTATGGGAAATTCCTCGATCGTACAGTGGTTACACTCGATCC 487  
Qy 377 LeuSerLysSerTTPileValProGlyTTPArgLeuGlyTTPValAlaValTyrAspPro 396  
Db 488 ATATCAAGGGATCGAAAGTCCAGGATGCGAACTGGTTGGTTCACGCTACATGATCTA 547  
Qy 397 ArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrSerThrLeuAsnValSer 416  
Db 548 GACGGTGTCTTCAGAACACCAAGCTCTTACAGCTCTCAAGATTCTTCCAGATAAC 607  
Qy 417 ThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGlu 436  
Db 608 AATAACCTCCGACAGATATCCAGCGGCTATTCTCGACATCTTGGAGAAAACCTCCTCAA 667  
Qy 437 AspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGln 456  
Db 668 GAGTTTGTGATAAGAGCGACAGTCTTCTGAAGATAAGTAGAATTGGTTATCTTAAG 727  
Qy 457 IleLysGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet 476  
Db 728 CTCAGTACATCTCTAGCTCCTACTTGCTACATGAACCCGAGGCTCGACCTCTTATGG 787  
Qy 477 ValLysLeuAsnLeuHisLeuLeuGluGluIleAspAspAspIleAspPheCysLys 496  
Db 788 ACCGAGCTTGATTATCGAGCTTTGTGACATCGAAGACCATCAAGACTTTTGCATAAAG 847  
Qy 497 LeuAlaLysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTTP 516  
Db 848 CTGTCTAAGAGAAAGAAACCTCGTCTGTATTACAGGGATTGCTTCAGTCAGAACACTGG 907  
Qy 517 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536  
Db 908 TTGAGGCATCTATCGATATGGAGACTCCGGTATTGGAGGATGATCGAAGAGATTGAAG 967  
Qy 537 SerPheCysGlnArg-----AsnLysLysArgAsnSerSerAsp 549  
Db 968 AGCTTCTCGATCGCCATTCCCAACAAAAGCTCCCTCCAAAGAC 1012  
  
RESULT 12  
AAC36514  
ID AAC36514 standard; DNA; 1184 BP.  
XX AAC36514;  
XX  
XX  
XX 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 14092.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 18-JUL-1999; 99US-0144086.  
PR 18-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.

```

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1,43e-53
Score: 783.50
Percent Similarity: 66.27%
Best Local Similarity: 45.07%
Query Match: 27.43%
DB: 21

US-10-019-783-2 (1-551) x AAC36514 (1-1184)

QY 217 SerAlaValAlaLuhisLeuSerGlnGlyValProTyrMetLeuSerAlaAspVal 236
Db 11 AGTGGCGGTAGCAGATGTTCTTAAACCAAGGTCTTCCAAAGAGGTTCACGACATGACGTG 70
QY 237 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThr 256
Db 71 TTTATGACCTTGGATGCCAACCAAGTATTGAGTCGCGGTAGACATTCTCGCTAAACCG 130
QY 257 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 276
Db 131 AAA---GCCAACCGTTTGTCTCCGAGTCCCGGTCCCATGGGACCTAGTCGCTCCATC 187
QY 277 PheAsnArgLeuGluValAlaArgHisPheAspLeuIleProAspGlyGlyTrpGluIleAsp 296
Db 188 TACAGAACCTTGGATGCCCGCCACTATAATTTCTCCAGAAAGAACCTTGAATCGAC 247
QY 297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 316
Db 248 TTTGATAGCGTCCGAGCGCTCGTGGACGAGAACACACATTTGCCATATTTATATCAACCCC 307
QY 317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336
Db 308 CACACCCCAAGTAAACACCTACTCCGAGGTCTATCTCAAAACGCTCGCTGAATGGGT 367
QY 337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
Db 368 AAGGAACCTCAAGATTATGTTGTTCTGACGAGGTTTTTAGATGGACATCTTTGGGAGT 427
QY 357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySer 376
Db 428 AACCCCTTTTGTCCCTATGGGAAAATTCGTGATCGTACCAAGTGTTCACATCGGATCC 487
QY 377 LeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspPro 396
Db 488 ATATCAAGGGATGGAAGTCCAGATGGCGAACTGGTTGGCTCAGCTACATGATCTA 547
QY 397 ArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 416
Db 548 GACGGGTCTTCCAGAAACCAACCAAGGTCTTACAAAGTGTCTCAGGATTTTCTCCAGATAAC 607

```

Qy	417	ThrAspProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGlu	436	PR	14-MAY-1999;	99US-0134218
Db	608	ATAAACCTCCGACAGTATCCAGCGCTATCTCGACATCTTGGAGAAACTCTCAA	667	PR	14-MAY-1999;	99US-0134219
Qy	437	AspPhePheIysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGln	456	PR	14-MAY-1999;	99US-0134221
Db	668	GAGTCTTCGATAAGAGCGCAGAGTTTTCTGAAGATAAAGTAGAATTGGTTATCTAAG	727	PR	14-MAY-1999;	99US-0134370
Qy	457	IleLeuGluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMet	476	PR	18-MAY-1999;	99US-0134768
Db	728	CTCAGTACATCTCAGCTCCTACTGCTACATGAACCCGAGCGCTGCACCTCTTATGG	787	PR	19-MAY-1999;	99US-0134941
Qy	477	VallyLeuAsnLeuHiSleuLeuGluGluIleAspAspIleAspPheCysCysLys	496	PR	20-MAY-1999;	99US-0135124
Db	788	ACCGAGCTTCATTATCGAGCTTTGTGGACATCGAAGACGATCAAGACTTTTGAATAAG	847	PR	21-MAY-1999;	99US-0135353
Qy	497	LeuAlaIysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyr	516	PR	24-MAY-1999;	99US-0136021
Db	848	CTTGCTTAAGAGAAACCTCGTCTGTTTTACCGAGGATTCATTCAGTCAGAAAGTTCG	907	PR	25-MAY-1999;	99US-0136392
Qy	517	ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys	536	PR	28-MAY-1999;	99US-0136782
Db	908	TTGAGCGATTCTATCGATATGGAGACTCCGGTATTGGAGGATTCATTCGAAAGTTCG	967	PR	01-JUN-1999;	99US-0137232
Qy	537	SerPheCysGlnArg-----AsnLysLysArgAsnSerSerAsp	549	PR	03-JUN-1999;	99US-0137528
Db	968	AGCTTCTGCGATCGCATTCACCAAAAAAGCTCCCTCCAAAGAC	1012	PR	04-JUN-1999;	99US-0137502
RESULT	13			PR	07-JUN-1999;	99US-0137724
ID	AAC45003			PR	08-JUN-1999;	99US-0138094
AC	AAC45003;			PR	10-JUN-1999;	99US-0138540
XX	XX			PR	10-JUN-1999;	99US-0138847
XX	XX			PR	14-JUN-1999;	99US-0139119
XX	XX			PR	16-JUN-1999;	99US-0139452
DE	18-OCT-2000 (first entry)			PR	16-JUN-1999;	99US-0139453
XX	XX			PR	17-JUN-1999;	99US-0139492
XX	XX			PR	18-JUN-1999;	99US-0139456
XX	XX			PR	18-JUN-1999;	99US-0139456
XX	XX			PR	18-JUN-1999;	99US-0139457
XX	XX			PR	18-JUN-1999;	99US-0139458
XX	XX			PR	18-JUN-1999;	99US-0139458
XX	XX			PR	18-JUN-1999;	99US-0139460
XX	XX			PR	18-JUN-1999;	99US-0139461
XX	XX			PR	18-JUN-1999;	99US-0139462
XX	XX			PR	18-JUN-1999;	99US-0139463
XX	XX			PR	18-JUN-1999;	99US-0139750
XX	XX			PR	18-JUN-1999;	99US-0139763
XX	XX			PR	21-JUN-1999;	99US-0139817
XX	XX			PR	22-JUN-1999;	99US-0139899
XX	XX			PR	23-JUN-1999;	99US-0140353
XX	XX			PR	23-JUN-1999;	99US-0140354
XX	XX			PR	24-JUN-1999;	99US-0140695
XX	XX			PR	28-JUN-1999;	99US-0140823
XX	XX			PR	29-JUN-1999;	99US-0140991
XX	XX			PR	30-JUN-1999;	99US-0141287
XX	XX			PR	01-JUL-1999;	99US-0141842
XX	XX			PR	01-JUL-1999;	99US-0142154
XX	XX			PR	02-JUL-1999;	99US-0142055
XX	XX			PR	06-JUL-1999;	99US-0142390
XX	XX			PR	08-JUL-1999;	99US-0142803
XX	XX			PR	09-JUL-1999;	99US-0142920
XX	XX			PR	12-JUL-1999;	99US-0142977
XX	XX			PR	13-JUL-1999;	99US-0143542
XX	XX			PR	14-JUL-1999;	99US-0143624
XX	XX			PR	15-JUL-1999;	99US-0144005
XX	XX			PR	16-JUL-1999;	99US-0144085
XX	XX			PR	16-JUL-1999;	99US-0144086
XX	XX			PR	19-JUL-1999;	99US-0144325
XX	XX			PR	19-JUL-1999;	99US-0144331
XX	XX			PR	19-JUL-1999;	99US-0144332
XX	XX			PR	19-JUL-1999;	99US-0144333
XX						



```
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0145388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158234.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 5,24e-53
Score: 778.50
Percent Similarity: 60.95%
Best Local Similarity: 40.95%
Query Match: 27.26%
DB: 21
Gaps: 9

US-10-019-783-2 (1-551) x AAC45003 (1-1584)

Qy 130 ValGluTrpAsnPheAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsnMet 149
Db 88 GTAGATTGGCAGTTTCAGTGGCAGTGAT-----GCACCGCAAAAGGCCGCCCAAGCC 138
Qy 150 SerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArgPro 169
Db 139 TCCTAGGCACCTTACTCTTCTGAAATCTTCAGTCTGTGTGACCTCAAGA---AAGCCT 195
Qy 170 ValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaValGlu 189
Db 196 ATTTTG-----CCTCCATAAGTGAAGAGGACGAGACGAGCCATACC 237
Qy 190 AlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCysTyrProAlaGly 209
Db 238 GCGRAAAGCGGTTGTTAAGGCTTCCTTGGGTACGGGAACGCCCTAAGTCCAAGT 297
Qy 210 ValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValProTyr 229
Db 298 ATTGGCTCCCGGTGCCAAACGTCGTGGCGGAATATCTAAACCGAGATCTTGACAAT 357
Qy 230 MetLeuSerAlaAspAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluValIle 249
Db 358 AAGCTAACGGGAGATGACGTATATATGACTTTGGATGCAACAAACAAATCGAGCTCGCG 417
Qy 250 IleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeuProArgProGlyTyrPro 269
Db 418 GTTAGTATTCTGGCTAAACCGAAA---GCCAACATCTTGCTTCGAGGCGCAGGCTCCCA 474
Qy 270 AsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIlePro 289
Db 475 TGGACATGGTCCACTCTATCTACAAACACCTTGAGTCCGAAAGGTATGAGTTTCATCCT 534
Qy 290 AspLysGlyTrpGluIleAspIleAspSerLeuGluSerIleAlaAspLysAsnThrThr 309
Db 535 GAAAGAGACTTTGAGATCGATTTCACAGCGTTAGAGAGATGGTAGACAGACACTTTT 594
Qy 310 AlaMetValIleIleAsnProAsnAsnProCysGlySerValTyrSerTyrAspHisLeu 329
Db 595 GCGATATTATATCAACCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCCCAATCC 654
Qy 330 SerLys-----ValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAsp 346
Db 655 AAGCA-GGTATTACCTCGCTACATGGCAGCGGAACCTGGGATATGCTGTTTCGGAT 713
Qy 347 GluValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGly 366
Db 714 GAAGTATATAGATGGTCCGTTATTCGGAGTATCCATTGTTCCATGGGAAATCTCTCC 773
Qy 367 HisIleThrProValLeuSerIleGlySerLeuSerLysSerTyrIleValProGlyTyr 386
Db 774 TCGATAGTACCAGTATTACACTCGGTCCTCATATCCAAAGGATGGATTGTCAGGATGG 833
```

QY 387 ArgLeuGlyTTPValAlaValTyAspProArgLysIleLeuGlnGluThrLysIleSer 406  
DB 834 CGAAGGGCTGGCTGGCTTACAGATCTAAATGGTGTCTTTAGATCCACAAAGGTCCTCG 893  
QY 407 Thr-----SerIleThrAsnTyLeuAsnValSerThrAspProAlaThrPheIleGln 424  
DB 894 GTCTTAAAGTGGCAAGAAATTTCTGGAGATTACTTCCAAACCCCAACTGTCAATCCAG 953  
QY 425 AlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleIleGly 444  
DB 954 CGGCTATTCACAACTTTCGAGAACTCTCAAGATTTTTCGAGAAAGGGGATT 1013  
QY 445 LeuLeuLysGluSerSerGluIleCysTyrysglnIleLysGluAsnLysTyryIleThr 464  
DB 1014 TTTCTAAAGACAAAGTGGATTTTGGATATTTCTAAGCTCAAGACATACCAACTCTCACC 1073  
QY 465 CysProHisLysProGluGlySerMetPheValMet---ValLysLeuAsn---LeuHis 482  
DB 1074 TGCTACATCAACCTGAATCATGTGACCTTCTT-ATGAGACCAAGCTGGACCATTCAT 1132  
QY 483 LeuLeuGluGluIleAspAspIleAspPheCysCysLysLysLeuAlaLysGluLysSer 502  
DB 1133 TTTGTA---GACATTAAGACGATCAGCATCTCTGCAGGAAGCTTGTCTAAGGAAGAAAT 1189  
QY 503 ValIleLeuCysProGlySerValLeuGlyMetAlaSerTTPValArgIleThrPheAla 522  
DB 1190 CTCGTGTGTTTACAGGAATTCATTTGGTCAGATTAAGCTGTTGAGCGCAT 1249  
QY 523 CysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGlnArgAsn 542  
DB 1250 ATGGAGACTCCAGATTCGAGGATTCATTTGAAGATTAAGAGCTTCTGTGACGCCAT 1309  
RESULT 14  
ID ABN98840/c  
XX ID ABN98840 standard; DNA; 858 BP.  
AC ABN98840;  
XX  
XX  
DT 01-AUG-2002 (first entry)  
DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 608.  
XX  
XX Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;  
KW disease; crop; thale cress; tolerance factor; insect; pathogen;  
KW nutrition; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
XX US2002023281-A1.  
PN  
XX  
PD 21-FEB-2002.  
XX  
PF 26-JAN-2001; 2001US-0770445.  
XX  
PR 27-JAN-2000; 2000US-178472P.  
XX  
PA (GORL/) GORLACH J.  
PA (ANY/) AN Y.  
PA (HAM/) HAMILTON C M.  
PA (PRIC/) PRICE J L.  
PA (RAIN/) RAINES T M.  
PA (YUY/) YU Y.  
PA (RAME/) RAMEAKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHW A V.  
PA (LEDE/) LEDFORD B L.  
PA (WORS/) WOESSNER J P.  
PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
PA (KRIC/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.

PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
XX (HURB/) HURBAN P.  
XX  
PI Goriach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
XX  
XX MPI; 2002-400781/43.  
XX  
XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,  
PT producing compositions that modulate the expression or function of its  
PT encoded protein, and mapping functional regions of protein -  
XX  
XX Claim 1; SEQ ID NO 608; 49pp + Sequence Listing; English.  
XX  
XX The invention relates to an Arabidopsis thaliana nucleic acid (I)  
XX comprising a sequence capable of hybridising under stringent conditions  
XX to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),  
XX given in the specification or its fragment. A polypeptide (II) encoded by  
XX (I), a transgenic plant (III) comprising an exogenous nucleic acid or a  
XX genetically modified cell (IV) comprising an exogenous nucleic acid, is  
XX useful for screening a candidate agent for its biological effect. (I) is  
XX useful in identifying homologous or related genes, in producing  
XX compositions that modulate the expression or function of its encoded  
XX protein, mapping functional regions of the protein and in studying  
XX associated physiological pathways. (I) is also useful for the genetic  
XX manipulation of cells, particularly plant cells. (I) is also useful in  
XX screening assays of various plant strains to determine the strains that  
XX are best capable of withstanding a particular disease or environmental  
XX stress. (II) and (III) are useful for screening of biologically active  
XX agents, e.g. fungicides, insecticides, etc., for elucidating biochemical  
XX pathways. The screened agents are useful in improved methods of treating  
XX crops to prevent or treat disease. (II) are also useful in screening  
XX programs to identify agents that mimic or enhance the action of tolerance  
XX factors. Such agents are useful in improved methods of treating crops to  
XX enhance their tolerance to environmental stress. (I) is also useful  
XX for enhancing or inhibiting production of a biosynthetic product in a  
XX plant. (III) is useful for identifying other mediators that may induce  
XX expression of proteins of interest, for establishing the extent to which  
XX any specific insect and/or pathogen is responsible for damage to a  
XX particular plant, for identifying other mediators that enhance or induce  
XX tolerance to environmental stress, for identifying factors involved in  
XX biosynthetic pathways of nutritional, commercial, or medicinal value and  
XX for identifying productions of nutritional, commercial or medicinal  
XX value. (IV) is useful in the study of genetic function and regulation,  
XX for alteration of the cellular metabolism and for screening compounds  
XX that may affect the biological function of the gene or gene products.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from USPTO  
XX at seqdata.uspto.gov/sequence.html?DocID=999909770445.  
XX  
SQ Sequence 858 BP; 254 A; 188 C; 165 G; 251 T; 0 other;

Alignment Scores:  
Pred. No.: 4,348-44 Length: 858  
Score: 663.00 Matches: 119  
Percent Similarity: 70.87% Conservative: 44  
Best Local Similarity: 51.74% Mismatches: 67  
Query Match: 23.21% Indels: 0  
DB: 24 Gaps: 0

US-10-019-783-2 (1-551) x ABN98840 (1-858)

QY 320 CysGlySerValTySerTyAspHisLysLeuSerLysValAlaGluValAlaLysArgLeu 339  
DB 858 TGTGGGAATGTTTCTCTCGCCCAACATCTTCAAAAGATTGCAGACGCGCTTGCAGGCTT 799  
QY 340 GlyIleIleValIleAlaAspGluValTyryGlyLysLeuValLeuGlySerAlaProPhe 359  
DB 798 GGTATATCTTGTGATCCGACGAGGAGTCTATGACCATTTTGCCTTTGGGGATTAACCATTT 739

```

QY 360 ileProMetGlyValphecHisIleThrProValLeuSerIleGlySerLeuSerIlys 379
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 738 GGTCCATGGCAGAGTTCGACAGTAGTGCCTGTTATTTAGTCTAGGTCATATCTAA 679
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 380 SerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspProArgLysIle 399
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 678 AGATGGTTTGTCTCGATGGAGACTTGGTGGATGGTGACTCTTGACCTCATGGCATC 619
QY 400 LeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPro 419
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 618 ATGAAGATTTGGGTTTGGTTCAGACTCTTATCAATGTGTCAACATGTCACGAGTCT 559
QY 420 AlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePhe 439
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 558 GCAAGCTTTATTCAGGAGCAATGCTGATATCATTTGGGATACAAAGGAGAGTTCTTC 499
QY 440 LysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGlu 459
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 498 TCATCAAACTTGAATGGTGAATAATGTGCAGAGATTGTTATGAGGAGCTTATGAAG 439
QY 460 AsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeu 479
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 438 ATTCTTGCATCACTTGCCCTGCAACCCGAGGCTCAATGTTACAGTGTGTAAGTTA 379
QY 480 AsnLeuHisLeuLeuGluGluIleAspAspIleAspPheCysLysLeuAlaLys 499
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 378 AACTTTTCACTACTCGAAGATATCATGATGATTTGGACTTCTGCTCCAGCTGGCTAA 319
QY 500 GluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArgIle 519
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 318 GAGGAACTATGATCATCTTACCAGGTCAAGCTGTGGGCTTAAGAACTGGCTACGATC 259
QY 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys 539
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 258 ACCTTGCAGTTGAGCTTCTCATAGAGGTTTTCAGGCTAAGAACTTTACT 199
QY 540 GlnArgAsnLysLysArgAsnSerSerAsp 549
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 198 GAGAGACTCCAAAGAAATCAGCCATGAGAC 169

RESULT 15
ID ABQ76730 standard; cDNA; 1377 BP.
XX AC ABQ76730;
XX DT 03-MAR-2003 (first entry)
XX DE Rat tyrosine aminotransferase cDNA fragment.
XX KW Rat; tyrosine aminotransferase; Vitamin E; TAT; fat-soluble antioxidant;
KW food; feed; pharmaceutical; cosmetic; abiotic stress resistance; cold;
KW frost; drought; heat; salt; tocotrienol concentration; ss.
XX OS Rattus norvegicus.
XX FH Location/Qualifiers
XX FT 1..1371
XX FT /*tag= a
XX FT /product= "tyrosine aminotransferase"
XX FT /partial
XX PN DE10111676-A1.
XX PD 12-SEP-2002.
XX PF 09-MAR-2001; 2001DE-1011676.
XX PR 09-MAR-2001; 2001DE-1011676.
XX PA (SUNG-) SUNGENE GMBH & CO KGAA.
XX FI Badur R, Geiger M;

```

```

XX WPI; 2003-047848/05.
DR P-PSDB; ABG73585.
XX Preparation of vitamin E in genetically modified organisms, useful as
PT antioxidant in e.g. food or cosmetics, by overexpression of tyrosine
PT aminotransferase
XX Claim 5; Page 13-15; 24pp; German.
XX This invention describes a novel method for the preparation of Vitamin E
CC by growing organisms having greater tyrosine aminotransferase (TAT)
CC activity than the wild type. The invention describes a construct
CC containing a sequence encoding TAT, linked to at least one
CC transcriptional and translational control signal, which is used in the
CC production of genetically modified organisms (GMO) in which the
CC modification increases TAT activity relative to wild type. The method is
CC used to produce vitamin E by from the GMO for use as fat-soluble
CC antioxidant in food, feeds, pharmaceuticals and cosmetics, to produce
CC plants for use as food or feed, or supplements or to produce plants with
CC increased resistance to abiotic stress (cold, frost, drought, heat and
CC salt). Overexpression of TAT selectively increases the content of
CC tocotrienols relative to tocopherols. This sequence encodes the partial
CC rat TAT protein described in the disclosure of the invention.
XX SQ Sequence 1377 BP; 332 A; 350 C; 368 G; 327 T; 0 other;

```

## Alignment Scores:

```

Pred. No.: 2,666-43 Length: 1377
Score: 656.50 Matches: 140
Percent Similarity: 54.81% Conservative: 88
Best Local Similarity: 33.65% Mismatches: 165
Query Match: 22.99% Indels: 23
DB: 25 Gaps: 8

```

US-10-019-783-2 (1-551) x ABQ76730 (1-1377)

```

QY 147 AlaAsnMetSerIleArgAlaIleArgTyrLysIle---SerAlaSerValGlnGluLys 165
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 151 TCCATAAGACCTTCAATCCATCCAGCCATCGTGGACACATGATGAAGTGCAGCCCAAT 210
QY 166 GlyProArgProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArg 185
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 211 CGAACCAAGACCCGTGATTTCTCTGTCATTTGGGACCTTACTGTTGGGAACCTGCCT 270
QY 186 ThrAlaValGluAlaGluAspAlaValAlaAlaValArgThrGlyGlnPheAsnCys 205
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 271 ACAGACCTTGAGTTACCCAGCCATGAAGATGCCCTGACTCGGGAGTACATGCGC 330
QY 206 TyrProAlaGlyValGlyLeuProAlaAlaArgSerAlaValAlaGlu-----HisLeu 223
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 331 TATGCCCGCTCCATCGCTACCTATCCAGTCGGGAGGAGTCCCTTCTTACTACCACTGT 390
QY 224 SerGlnGlyValProTyrMetLeuSerAlaAspAspValPheLeuThrAlaGlyThr 243
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 391 CATGAGGCT-----CCTCTGGAAGCTAAGATGTCATCTGCAGAGCGGTGCACT 441
QY 244 GlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuLeu 263
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 442 CAGGCCATTCAGCTAGTCTAGCTGTGTTGGCC---AATCCTGGACAAACATCCTCATT 498
QY 264 ProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuValArg 283
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 499 CCAAGGCCCGGGTTTTCCCTCTATAGGACTTTGGGTGAGTCTATGGGAATGAGGTCAAG 558
QY 284 HisPheAspLeuIleProAspLysGlyTrpGluIleAspIleAspSerLeuGluSerIle 303
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 559 CTCACATCTCTCGCTGAGAGTCTTGGGAATGACCTAAACAACTGAATCTCTG 618
QY 304 AlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnProCysGlySerVal 323
Db   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 619 ATCGATGAAAAAACACGCGTGTCTTGTGTCAACACCCATCCATCCCTTGGCTCCGTTG 678

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 8, 2003, 17:15:50 ; Search time 6070.26 Seconds  
(without alignments)  
3713.386 Million cell updates/sec

Title: US-10-019-783-2

Perfect score: 2856

Sequence: 1 MATVRQSDGVAAVLAA.....LGRKSPQCRNKRNSDDC 551

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cn2\_1/USPTO.spool/US10019783/runat\_07112003\_161643\_19836/app\_query.fasta\_1.1358  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=sct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10019783 @CN 1.1.6944 arunat 07112003 161643 19836 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.man.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2856	100.0	1895	8	AB005788 Hordeum v
2	2484	87.0	10966	6	BD012010 Creation
3	2484	87.0	10966	6	E49893 Creation of
4	2484	87.0	10966	8	AB024006 Hordeum v
5	2142.5	75.0	1660	8	D88273 Hordeum vul
6	1219	42.7	1290	6	AX653727 Sequence
7	1126.5	39.4	1600	8	AY054204 Arabidops
8	1124.5	39.4	1611	8	AF301900 Arabidops
9	1123.5	39.3	1611	8	AF301898 Arabidops
10	1123.5	39.3	1611	8	AF301899 Arabidops
11	1123.5	39.3	1683	8	AY050987 Arabidops
12	1121.5	39.3	920	6	AX653224 Sequence
13	1121.5	39.3	1389	6	AX575731 Sequence
14	1119	39.2	1243	6	AX575733 Sequence
15	1119	39.2	1276	8	BT001912 Arabidops
16	1118.5	39.2	1389	6	AX507289 Sequence
17	1118.5	39.2	1420	8	AY031293 Arabidops
18	1118	39.1	1570	8	BT000782 Arabidops
19	1117	39.1	1233	6	AX654009 Sequence
20	1103	38.6	1236	8	CBL458993 Coleus bl
21	1097	38.4	1564	8	AY113848 Arabidops
22	1008.5	35.3	1263.5	2	AP004116 Oryza sat
23	1005.5	35.2	15980	2	AP005743 Sequence
24	998	34.9	1338	6	AX412376 Sequence
25	998	34.9	1338	6	AX507174 Sequence
26	998	34.9	1338	6	AX651743 Sequence
27	998	34.9	1338	8	BT006593 Arabidops
28	998	34.9	1495	8	BT002475 Arabidops
29	992	34.7	1535	8	AY085324 Arabidops
30	987	34.6	1334	6	AX575729 Sequence
31	965	33.8	93695	8	ATF2009 Arabidops
32	965	33.8	199749	8	ATCHRIV68 Arabidops
33	948.5	33.2	70475	8	AB018112 Arabidops
34	899.5	31.5	61384	8	AC007048 Arabidops
35	895	31.3	1201	8	AY142527 Arabidops
36	883.5	30.9	73840	8	AB007644 Arabidops
37	879	30.8	151073	2	AC134624 Oryza sat
38	877.5	30.7	930	6	AX653614 Sequence
39	872.5	30.5	1269	6	AX575727 Sequence
40	872.5	30.5	1462	8	AF268090 Arabidops
41	872.5	30.5	1495	8	BT000307 Arabidops
42	872.5	30.5	1634	8	AY099811 Arabidops
43	864	30.3	1284	8	AY187682 Brassica
44	836.5	29.3	1306	8	AY123024 Arabidops
45	836.5	29.3	1563	8	AY070389 Arabidops

# ALIGNMENTS

RESULT 1



```
Db      1216 TGGATAGTCCCTGGATGGCGGTTCGATGGGTAGCGGTGACGACCCGAGAAAGATCTTA 1275
Qy      401 GlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspProAla 420
Db      1276 CAGGAAACTTAAGATCTTACATCAATACGAAATACCTCAATGTCTCGACAGACCCAGCA 1335
Qy      421 ThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLys 440
Db      1336 ACCTTCATTCAGCGAGCTCTTCTCAGATCTTGAGAACACAAAGGAAGATTTCTTAAG 1395
Qy      441 AlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsn 460
Db      1396 GCGATTATTGGTCTGCTAAAGGAATCATCAGAGATATGCTACAAACAATAAAGGAAAC 1455
Qy      461 LysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeuAsn 480
Db      1456 AAATACATTCATACATGCTTCACAAAGCCAGAGGATCAATGTTTGTCTCATGTTGAACTGAAC 1515
Qy      481 LeuHisLeuLeuGluGluIleAspAspAspIleAspPheCysLysLeuAlaLysGlu 500
Db      1516 TTACATCTTTTGAGGAATAACAGATGACATGATGATTTTGTCTGCAAGCTCGCAAGAA 1575
Qy      501 GluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyrValArgIleThr 520
Db      1576 GAATCAGTAATCTTATGCCCAGGAGTGTCTTGGATGGCAAACTGGGTCCGCATTACT 1635
Qy      521 PheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCysGln 540
Db      1636 TTTGTTGTGTTCCATCTCTCTTCAAGATGCTCTCGAAGATCAATATCAITCTGTCAA 1695
Qy      541 ArgAsnLysLysArgAsnSerSerAspAspCys 551
Db      1696 AGGAACAAGAGAGAAATTCGAGCGATGATTGC 1728
```

```
RESULT 2
BD012010
LOCUS      BD012010          10966 bp      DNA      linear      PAT 02-AUG-2002
DEFINITION Creation of iron-deficiency resisting rice plant.
ACCESSION  BD012010
VERSION    BD012010.1  GI:22092199
KEYWORDS  WO 0101762-A/1.
SOURCE    unidentified
ORGANISM  unidentified
REFERENCE  1 (bases 1 to 10966)
AUTHORS   Mori, S., Nakanishi, H., Takahashi, M. and Nishizawa, N.
TITLE     Creation of iron-deficiency resisting rice plant
JOURNAL   Patent: WO 0101762-A 1 11-JAN-2001;
          JAPAN SCIENCE AND TECHNOLOGY CORP, SATOSHI MORI, HIROMI NAKANISHI,
          MICHIKO TAKAHASHI, NAKO NISHIZAWA
COMMENT   OS  Hordeum vulgare L. var. Igri
          PN  WO 0101762-A/1
          PD  11-JAN-2001
          PF  04-JUL-2000 WO 2000JP004425
          PR  05-JUL-1999 JP 99P 190318
          PI  SATOSHI MORI, HIROMI NAKANISHI, MICHIKO TAKAHASHI, PI  NAKO
          NISHIZAWA
          PC  A01H5/00, C12N5/14, C12N15/52
          CC
```

```
FEATURES
             Location/Qualifiers
             1..10966
                /organism="unidentified"
                /mol_type="genomic DNA"
                /db_xref="taxon:32644"
BASE COUNT  3082 a 2422 c 2366 g 3096 t
ORIGIN
Alignment Scores:
Pred. No.:      1.47e-131      Length:      10966
Score:          2484.00        Matches:      550
Percent Similarity: 51.30%      Conservative: 1
```

```
Best Local Similarity: 51.21%      Mismatches: 0
Query Match:          86.97%      Indels: 523
DB:                   6          Gaps: 6
US-10-019-783-2 (1-551) x BD012010 (1-10966)
Qy      1 MetAlaThrValArgGlnSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAlaA 20
Db      650 ATGCCCAACCTAGCCAGAGCGAGTCCGCCGACGAGCGCTTGCCTGGCGCGAGCC 709
Qy      21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAlaValAsnGlyLysSerAsnGly 40
Db      710 GCGAACGCGCAAGAGCAACGCCATGCGTGGCTGCCGCGTGAACGCGAAGACGACGC 769
Qy      41 HisGlyValAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60
Db      770 CATGGCGTGGATCCGACGCGAACGCGAAGACGCGCATCGCGTGGCTGGCGAGCGG 829
Qy      61 AsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAsnGlyHisGlyGluAlaThrAla 80
Db      830 ACGGCAAGAGCAACGCCATGCGAGGCCACTGCGAACGCGCCACGCGAGGCGCATTCGG 889
Qy      81 AsnGlyLysThrAsnGlyHisArgGluSerAsnGlyHisAlaGluAlaAlaAspAlaAsn 100
Db      890 AACGCAAGACCAACCGCCACCGCGAGAGCAACGCGCATGCTGAGCGCGCGACGGAAC 949
Qy      101 GlyLysSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyLysSerAsnGlyHisAla 120
Db      950 GCGAGAGCAACAGACATGCCAGGACTCCCGCGCAACGCGGAGAGCAACGCGCATCG 1009
Qy      121 AlaAlaAlaGluGluGluAlaValGluTyrAsnPheAlaGlyAlaLysAspGly 140
Db      1010 GCGGCGCGCGCAGAGGAGGAGGCGGTGGAGTTCGGAATTTCCGCGGTGCCAAGGACGC 1069
Qy      141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160
Db      1070 GTGCTGCGCGCGAGCGGCGGCAACATGAGCATCCGCGCGATACGGTACAGATCAGCGCG 1129
Qy      161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
Db      1130 AGCGTGCGAGGAGAGGCGCGCGCGTCTCGCTGGCCACCGCGGAGCCCGTCCGTG 1189
Qy      181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArgThr 200
Db      1190 TTCCCGGCTTCGCGACGCGCGTGCAGCGCGAGAGCGCGTCCGCCGCGCGTGCACCC 1249
Qy      201 GlyLysPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArg 216
Db      1250 GCGCAGTTTCACTGCTACCCCGCGCGGTTCGGCTTCGCCCGCCACGAGGTAAACAACAA 1309
Qy      216 ----- 216
Db      1310 CAACAACACAGAACAACATTTCTTTTCGGGTGTCGTGTCGGGCGCAATCCATGTCATCG 1369
Qy      216 ----- 216
Db      1370 CATGTGCGCGTTTCACGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1429
Qy      217 -----SerAlaVal 219
Db      1430 ATGAGAAATCTGACCTTCTCCACCTTATACCAACAACAAACAAACAAACACACGCGCGTG 1489
Qy      220 AlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspAspValPheLeuThr 239
Db      1490 GCAGAGCACCTGTCGAGGCGGTGCGGTATCATGCTATCGCGCGACGACGCTTCTCTCACC 1549
Qy      240 AlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGlyVal 259
Db      1550 GCGGCGGGNACCCAGCGGATCGAGGTCAATAATCCCGGTGCTGCGCCAGACCGCGCGGCC 1609
Qy      260 AsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArg 279
Db      1610 AACATTCTGCTCCCGAGGCGAGGCTACCCAAACTACGAGGCGCGCGCGCGCTTCAACAGG 1669
```

```
QY 280 LeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSer 299
Db 1670 CTGGAGGTCGGGCATTTTCGACTCATCCCGACAGGGTGGGAGATCGACATCGACTCG 1729
QY 300 LeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnPro 319
Db 1730 CTGGATCCATCGCCGACAGACACACCGCCATGGTTCATCATAAACCCCAACAACCCG 1789
QY 320 CysGlySerValTyrSerTyrAspHisLeuSerLys 331
Db 1790 TCGCGACGGTTTACTCTACGACCATCTGTCCAAGGTTTCACATCCCTTGGCCTTGTCTGA 1849
QY 331 331
Db 1850 ATATGGATTTCAGTTCAGTGCACCTGCTGAATCTTTTGGCCAATCGCATCTGACTGATG 1909
QY 332 ValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGlu 347
Db 1910 TTGCTCAATTAGTTCGCGAGGTGGGAAAGGCTCGGATATTTGGTGATTTGCTGACGAG 1969
QY 348 ValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHis 367
Db 1970 GTATACGGCAAGCTGGTCTGGGACGCCGCCGTTTCATCCCAATGGGAGTGTGGGCAC 2029
QY 368 IleThrProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArg 387
Db 2030 ATCACCCCTGTGCTCTCCATAGGCTCTGTCCAAGTCATGGATAGTGCCTGGATGGCGG 2089
QY 388 LeuGlyTyrValAlaValTyrAspProArgLysIleLeuGlnGluThrLys 404
Db 2090 CTGTGATGGTAGCGGTGTACGCCGCCGAAAGATCTTTACAGGAACTAAGGTACTTAAA 2149
QY 404 404
Db 2150 TCCTATATCATCTTTTCAAAATGCTACTAAGGTGATTAATTAGTACTACTGTACATAT 2209
QY 405 IleSerThrSerIleThrAsnTyrLe 413
Db 2210 ATTTGCTAAATTTGACTGACATTTTGTGTGTAGATCTCTACATCAATTACGAATTAACCT 2269
QY 413 uAsnValSerThrAspProAlaThrPheIle 423
Db 2270 CAATGCTCGACGACCCAGCAACCTTCATTAGTGTAGTCTTTGGTATTTTACCTCGTTT 2329
QY 423 423
Db 2330 CAAGAATAAAGTCTTTGGTATTTACTCTCTCTTGTCTTATTTTGTCTCCGGTCCCTATGT 2389
QY 423 423
Db 2390 TGTAGGACCCCAAGTGCATGTCAAGTACCGTTTTTTTCACATTAAGTTTGAAGTCAAA 2449
QY 423 423
Db 2450 GTACAGACATACACTTGTAGTTATTTTACCTTTTGTCTTGTATCGATAAAATAAAA 2509
QY 423 423
Db 2510 AAATAAATAAAGTGAACCTGCTGTGTAATAAACCACTGTTCTTACAGATATACATGAT 2569
QY 423 423
Db 2570 TCACATATGGGCATCCCATATTTCTTTGGGTCAAGTATGACGATGTTTGAACCTCTTTT 2629
QY 423 423
Db 2630 AGAAAAATAGATACATTGTACTATGATATACCAATTTTATTAAAGAAATTTTCATATTTGATA 2689
QY 423 423
Db 2690 TCCTTGATGGTATTGTTCTCTGTGATTTCACACGATTTACTTGTGGTTTTTTGTACTATC 2749
```

```
424 -----GlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePheLysA 441
Db 2750 AAATTTGTTTCAGGAGCTCTTCTCAGATTCCTGAGACACAAGAGAGATTTCTTTAAGG 2809
QY 441 laIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGluAsnL 461
Db 2810 CGATTTATTGCTGTCTAAAGGAATCATCAGAGATATGCTACAAACAAATAAAGGAAACA 2869
QY 461 ystTyrIleThrCysProHisLysProGluGlySerMetPheValMet 476
Db 2870 AATACATTACATGCTCTCAAGCCAGAGAGATCAATGTTTGTCTATGTAAGCCTATTTT 2929
QY 476 476
Db 2930 GTGAAGTAAAAAATCTTAGGGAGTGTCAAGTAATCATAAACTTATTATATAGGATTAAT 2989
QY 476 476
Db 2990 CTGGGACCGAAATGCATCCAAATAATTAATTTCAAAATTCAAATTCAAATTCATTTCTCC 3049
QY 476 476
Db 3050 GTACATATTTTGAAGATGCATGATTTTAAAGAATAATGACGAGAGCTAAAGTTATGCTA 3109
QY 476 476
Db 3110 CGACTAATCATCTGGATATCTTTTGTCCATCTTTTGTATATACTGTGAATGTTAATGTT 3169
QY 476 476
Db 3170 CAAATCATATTACACAATAATCCATGCTAGTTTCTAGAAAGATGATTTTCTTTCTGTA 3229
QY 477 -----ValLysLeuAsnLeuHisLeuLeuGlu 485
Db 3230 CCATGAACCTCCGTATTAACTTCCATGTAAACAGAGTCAAACTGAACCTTACATCTTTTGGAG 3289
QY 486 GluIleAspAspIleAspPheCysCysLysLeuAlaLysGluGluSerValIleLeu 505
Db 3290 GAAATAGACATGACATTTGCTCAAGCTCGCAAAAGAAAGATTCAGTAATCTTA 3349
QY 506 CysPro 507
Db 3350 TGCCAGGTAGGAATCCATTGTTGATTTTGTACTGTATATGAAGTCTTATCAATTCOG 3409
QY 507 507
Db 3410 AGATGACTATACATATAATGATTAATCCATATTATGTCAGAAATTTGTATAACAGTGTAG 3469
QY 507 507
Db 3470 AATATTCTGGAAGACTTTTTTAAACAATAATTTCTGTGAAGACTAGATATCATGTACTTC 3529
QY 507 507
Db 3530 TCCTTGTCTTGTGACCTGATGTCCTTGTGCATGTTTGTCTCTCAAAAAAATAGC 3589
QY 507 507
Db 3590 AAGCACATGTTTCAATAATTTGTTAATAATAAATTTAGCCTTTTATATATGTTCTA 3649
QY 507 507
Db 3650 TTTTGAGATATTTTGTAGTCCAACTTATATATTTTGTGACTATTTCTCAAAAAAACAACCT 3709
QY 508 -----GlySerValIleGlyMetAlaAsnTrpValArgIle 519
Db 3710 ATATATGTGTGCTCTCAAAATGATAGGAGTGTCTTGGATGCGCAACTGGTCCGCATT 3769
QY 520 ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys 539
Db 3770 ACTTTTGTGTGTTCTCTCTTCAAGTGTCTCGGAAGGATCAAAATCATCTCTGT 3829
QY 540 GlnArgAsnLysLysArgAsnSerSerAspCys 551
```



```

Db      3830 CAAGGACAGAGAGAAATTCGAGGATGATTC 3865
|||||
RESULT 3
LOCUS   E49893
DEFINITION Creation of iron-deficiency resisting rice plant.
ACCESSION E49893
VERSION   E49893.1 GI:18629372
KEYWORDS JP 2001017012-A/1.
SOURCE   unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 10966)
AUTHORS Mori, S., Nakanishi, N., Takahashi, M. and Nishizawa, N.
TITLE     Creation of iron-deficiency resisting rice plant
JOURNAL   Patent: JP 2001017012-A 1 23-JAN-2001;
          SCIENCE & TECH AGENCY, SATOSHI MORI
COMMENT   OS Hordeum vulgare L. var. Igri
          PN JP 2001017012-A/1
          PD 23-JAN-2001
          PF 05-JUL-1999 JP 1999190318
          PR
          PI SATOSHI MORI, NOBUHITO NAKANISHI, MICHIO TAKAHASHI, PI NAKO
          NISHIZAWA
          PC A01H5/00, C12N5/10, C12N15/09//C12N9/10, C12N5/00, C12N15/00 CC
          FH Key
          FT source
          FT Location/Qualifiers
          FT 1. .10966
          FT /organism='Hordeum vulgare L. var. Igri'.
          FT Location/Qualifiers
          FT 1. .10966
          FT /organism='unidentified'
          FT /mol_type='genomic DNA'
          FT /db_xref='taxon:32644'
BASE COUNT 3082 a 2422 c 2366 g 3096 t
ORIGIN
Alignment Scores:
Pred. No.: 1.47e-131 Length: 10966
Score: 2484.00 Matches: 550
Percent Similarity: 51.30% Conservative: 1
Best Local Similarity: 51.21% Mismatches: 0
Query Match: 86.97% Indels: 523
DB: 6 Gaps: 6
US-10-019-783-2 (1-551) x E49893 (1-10966)
Qy      1 MetAlaThrValArgGlnSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20
Db      650 ATGGCCACCGTACGCAGAGCAGCAGGAGTCGCGCGGACGCGCTTGGCGTGGCCGCGC 709
Qy      21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAlaValAsnGlyLysSerAsnGly 40
Db      710 GCGAACGGCAGAGCAACGGCCATGGCGTGGTGGCGCGCGTCAACGGCAAGCAACGGC 769
Qy      41 HisGlyValAlaAspAlaAspAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAspAla 60
Db      770 CATGGCGTGGTGGCGAGCGCAACGGCAGAGCAACGGCCATGGCGTGGCGTGGCGAGCGC 829
Qy      61 AsnGlyLysSerAsnGlyHisAlaGluAlaThrAlaAlaAsnGlyHisGlyGluAlaThrAla 80
Db      830 AACGCGAAGAGCAACGGCCATCCCGAGGCCATCGCGAAGCGCCACGGCGGCGCACTGCG 889
Qy      81 AsnGlyLysThrAsnGlyHisArgGlnSerAsnGlyHisAlaGluAlaAlaAspAlaAsn 100
Db      890 AACGCGAAGCAACCGCCACCGCAGAGCAACGGCCATGGTGGAGCGCGCGAGCGCGCAAC 949
Qy      101 GlyGluSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGlyHisAla 120
Db      950 GCGGAGAGCAACAGCATGCCGAGGACTCCGCGCGGCGAAGCGGAGAGCAACGGGCATGCG 1009
Qy      121 AlaAlaAlaAlaGluGluGluAlaValGluThrAsnPheAlaGlyAlaLysAspGly 140
|||||

```

```

Db      1010 GCGCGCGCGCAGAGGAGGAGGCGGTGGAGTGGAAATTCGCGGGTGCACAGGACGCG 1069
Qy      141 ValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAla 160
Db      1070 GTGTGCGCGCGCAGCGGGGCGAAACATGAGCATCGGCGGATACGGTACAAAGATCAGCGCG 1129
Qy      161 SerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSerVal 180
Db      1130 AGCGTGAGGAGAGGAGGCGCGCGCGTGTCCGCTGCGCCACCGGGACCCGCTCCGTG 1189
Qy      181 PheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaAlaArgThr 200
Db      1190 TTCCCGGCGCTTCGCGACGGCGCGTGGAGCGGAGCGCGCTCCGCGCGCGCTCGCGACC 1249
Qy      201 GlyGlnPheAsnGlyProAlaGlyValGlyLeuProAlaAlaArg----- 216
Db      1250 GGCAGTTCAACTGCTACCCCGCGCGTCCGCTCCCGCGCGCAGGAGTAAACAACAA 1309
Qy      216 ----- 216
Db      1310 CAACAACAACAACAATTTCCCTTTTCGCGTGTGCTGCGCGGCAATCCATGCATGCG 1369
Qy      216 ----- 216
Db      1370 CATGTGCGGCTTTCACTGCTCGTCCGTCGCGTCCCGTTCCTTCTCTCTCTCTCTCT 1429
Qy      217 -----SerAlaVal 219
Db      1430 ATGGAATCTGACCTTCTCCACCTTATACCAACAACAACAACAACAACAACAACAACA 1489
Qy      220 AlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThr 239
Db      1490 GCAGAGCACCTGTGCGAGGCGTGGCGTACATGCTATGCGCGCAGCAGCTTCTCTCAC 1549
Qy      240 AlaGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAla 259
Db      1550 GCGCGGCGGACCCCGCGATCGAGGTCATAATCCGCTGTGCGCGCAGACCGCGCGCGCC 1609
Qy      260 AsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsnArg 279
Db      1610 AACATTCTGTCTCCCGAGCGAGCTACCCAACTACGAGCGCGCGCGCTTCAACAGG 1669
Qy      280 LeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAspIleAspSer 299
Db      1670 CTGAGGTCCGCGATTTTCGACCTCATCCCGCAGAGGGTGGAGATCGACATCGACTCG 1729
Qy      300 LeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnPro 319
Db      1730 CTGGAATCCATCCCGCAGCAAGAACACACCGCGCATGGTCATCATAAACCCCAACACCG 1789
Qy      320 CysGlySerValTyrSerTyrAspHisLeuSerLys----- 331
Db      1790 TGGCGCAGCGCTTACTCTACGACCATCTGTCCAAGGTTTTCACATCTCTTTCCTTGTCTGA 1849
Qy      331 ----- 331
Db      1850 ATATGGAATTCAGTTCAGTCACCTGTGAATCTTTTTCGCAATCGCATCTACTGACTGATG 1909
Qy      332 -----ValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGlu 347
Db      1910 TTGCTCAATTAGTTCGCGAGGTGGGAAAGGCTCGGAATATTGGTGAATTGTGTGACGAG 1969
Qy      348 ValTyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHis 367
Db      1970 GTATACGGCAAGCTGTTCTGGCGACGCGCGCTTCATCCCAATGGAGTGTGGGCGAC 2029
Qy      368 IleThrProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrTrpArg 387
Db      2030 ATACCCCTGTGTGTCCATAGAGTCTCTGTCCAGTCATGGATAGTGTCTGTATGGCGG 2089
Qy      388 LeuGlyTrpValAlaValTyrAspProArgLysIleLeuGlnGluThrLys----- 404
Db      2090 CTTGGATGGGTAGCGGTGTACGACCCCGCAAGATCTTACAGGAAACTTAAGTACTATAA 2149

```

QY	404	-----	404
Db	2150	TCTCTATATCATCTTTTCAAATGCTACTAAGTGATTAAATTAGTACTACTGTACAATAT	2209
QY	405	-----	405
Db	2210	ATTGCTAAATTGTTACTGACATTTTGTGGTAGATCTCTACATCAATTAGCAATTACCT	2269
QY	413	uAsnValSerThrAspProAlaThrPheIle-       	423
Db	2270	CAATGCTCGACAGACCAGCAACCTTCATTACAGTCAGTCCTTTGGTATTTTACCTCGTTT	2329
QY	423	-----	423
Db	2330	CAAGAAATAAAGTCTTTGGTATTACTCCTCCTGTGCTATTTTGTCTCCGGTCCCTATGT	2389
QY	423	-----	423
Db	2390	TGTAGCAGCCCACTGTCATGTCAGTCACCGTTTTTTCACATTAAAGTTTGAAAGTCAAA	2449
QY	423	-----	423
Db	2450	GTCAGACATACACTGTGTAGTTATTACCTTGTGTGCTTGATCCGATAAAATAAAA	2509
QY	423	-----	423
Db	2510	AAATACAAAACTGAACCTACTGTTGAATATAACCACTGTTCTTTACAAGATATACATGAT	2569
QY	423	-----	423
Db	2570	TGCACTATGGCATGCCATATCTTTTGGGTCAAGTAGCAGTATGTTGAAOCTCTTTT	2629
QY	423	-----	423
Db	2630	AGAAAATAGATACATTTGCTATCAGTATACCACTTTTATTAAAGAAATTCATATTTTGATA	2689
QY	423	-----	423
Db	2690	TCCTTGATGTPATGTCTCTGTGATTCACACGATTACTGTGGTTTTTTGTPACTATC	2749
QY	424	-----	424
Db	2750	AAATTGTTCAGGCAGCTTCTCTCAGATCTTGTGAAACACAAAGAGAATTTCTTTAAGG	2809
QY	441	lallelleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGluIleLysGluAsnL	461
Db	2810	CGATTATGTGTCCTAAGGAATCATCAGAGATATGCTACAAACAATAAAGGAAAAACA	2869
QY	461	YstYrIleThrCysProHisLysProGluGlySerMetPheValMet-       	476
Db	2870	AATACATTACATGTCCTCACAGCCAGAGGATCAATGTTTGTTCATGTAGCCCTATTTT	2929
QY	476	-----	476
Db	2930	GTGAAGTAAAAAATCTTAGGGAGTGTCAAGTAATCATATAAACTTATTTATAGGATTAA	2989
QY	476	-----	476
Db	2990	CTGGACCGAAATGCATCCACATAATTACTTCAAATTCAAATTCAAATTACATTTCTTCC	3049
QY	476	-----	476
Db	3050	GTACATATTTTGAAGATGCATGTATTTTAAGATAATAGCAGAGCTAAAGTTATGCTA	3109
QY	476	-----	476
Db	3110	CGACTAATCATCTGGATATCCTTTGTCCATCTTTTGTATTATATCTGTGGAATGTTAATGGT	3169
QY	476	-----	476
Db	3170	CAAAATCATATTACAAATATCCATGCTAGTTTCTAGAAAAGATTGATTATTTTCTGTAA	3229

Qy	477	-----ValysLeuasnleuHisLeuLeuGlu	485
Db	3230	CCATGAACCTCCGTATTAAACTTCCATGTTAAACAGAGGTGAAACTGAACCTACATCTTTTGGAG	3289
Qy	486	GluLeaspAspAspIleAspPheCysCysLysLeuAlaLysGluGluSerValIleLeu	505
Db	3290	GAATATAGACCATGACATTTGTTTCTGCGAGCTCGCAAAAGAGAAATCAGTAATCTTTA	3349
Qy	506	CysPro-----	507
Db	3350	TGCCCAGGTAGGAATCCATTTGTTGATTTTTTGACTGTATATGAAGTTCTTTATCAATTTCCG	3409
Qy	507	-----	507
Db	3410	AGATGACTATACATATAAATGATACATATTATGTGCAGAAATGTATTAACAGTGTTAG	3469
Qy	507	-----	507
Db	3470	AATATTCTGGAAGACTTTTTTTAAACAATATTCTGTGAAGACTAGATATCATGTACTTC	3529
Qy	507	-----	507
Db	3530	TCCTTGTGTTTCTTGACCTGATGTCCTTGGTCACATGTTGTGTCCTCACAAAAAATAGC	3589
Qy	507	-----	507
Db	3590	AAGCACAATGTTTCAATAATTTGTTAATAATTAATTTAGCCTTTAATTTATATGGTTCTA	3649
Qy	507	-----	507
Db	3650	TTTTGAGATATTTTTTGTAGTCCAACTATATATATTTTGTGACTATTCTCAAAAAACAACATT	3709
Qy	508	-----GlySerValLeuGlyMetAlaAsnTrpValAlaGlyLe	519
Db	3710	ATATATGTGTGCTCTCAATGTAGGAGTGTTCTTTGGAAATGCGAACTGGGTCCCATTT	3769
Qy	520	ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys	539
Db	3770	ACTTTTGTTGTGTTCCATCTCTCTTCAAGATGTTCTCGAAGAGTCAAAATCATTTCTGT	3829
Qy	540	GlnArgasnlylslysArgAsnSerSerAspAspCys	551
Db	3830	CAAAAGCAAGAAGAGAAATTCGAGCATGATGC	3865
RESULT	4		
LOCUS	AB024006	10966 bp	DNA linear
DEFINITION	Hordeum vulgare naat-B and naat-A genes for nicotianamine aminotransferase, complete cds.		
ACCESSION	AB024006		
VERSION	AB024006.1	GI:6469088	
KEYWORDS	nicotianamine aminotransferase.		
SOURCE	Hordeum vulgare subsp. vulgare		
ORGANISM	Hordeum vulgare subsp. vulgare		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Hordeum.		
REFERENCE	1 (sites)		
AUTHORS	Takahashi,M., Yamaguchi,H., Nakanishi,H., Shioiri,T., Nishizawa,N.K. and Mori,S.		
TITLE	Cloning two genes for nicotianamine aminotransferase, a critical enzyme in iron acquisition (Strategy II) in graminaceous plants		
JOURNAL	Plant Physiol.		
MEDLINE	121 (3), 947-956 (1999)		
PUBMED	20027437		
REFERENCE	2 (bases 1 to 10966)		
AUTHORS	Mori,S., Takahashi,M. and Nakanishi,H.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-FEB-1999)		
MEDLINE	Satoshi Mori, University of Tokyo, Dept. of Applied Biological Chemistry; Rayoi 1-1, Bunkyo-ku, Tokyo 113-8657, Japan (E-mail:asmori@hongo.ecc.u-tokyo.ac.jp,		
FEATURES	Tel:81-3-3812-2111(ex.5106), Fax:81-3-3812-0544)		
	Tel:81-3-3812-2111(ex.5106), Fax:81-3-3812-0544)		
	Location/Qualifiers		

```

source
1. 10966
/organism="Hordium vulgare subsp. vulgare"
/mol_type="genomic DNA"
/sub_species="Igr1"
/db_xref="taxon:112509"
/clone_lib="lambda FIX11"
join(650..1299,1483..1825,1922..2140,2244..2303,
2761..2916,3263..3356,3735..3868)
/gene="naat-B"
CDS
join(650..1299,1483..1825,1922..2140,2244..2303,
2761..2916,3263..3356,3735..3868)
/gene="naat-B"
/codon_start=1
/product="nicotianamine aminotransferase"
/protein_id="BAAB7054.1"
/db_xref="GI:6469089"
/translation="MATVRQSDGVAAVLAALAAANGKSNHGHHVAAVANGKSNHGVD
ADANGKSNHGVAADANGKSNHAEATANGHEATANGKSNHRESNGHAAADANGE
SNEHAESAANGENGAHAAAEBAEVEWNAFAGKDGVLAAATGANNYSIRAIRYKISA
SVQKGPRLVPLAHGDPSPVFAPTAVEAEDAAALRTGQFNCPAGVGLPAARSA
VAEHLSDGVPVYMLSADVDVLTAGTQALEVILPVLAOTAGANILLPSPGYPNVEARAA
ENREVRHFDLIPKGMHEIDDSLESADKNTTAMVILNPNPFGSVYSDHLSKVAE
VAKRLGILVADENVYKLVLSGAPFIEMGVFHHITPVLSIGLSKSMIVFGWLGWVA
VYDRPKLQETKISTINYLNVDTPATFIOAALPOLLENTKEDFFPAIIGLLKRES
EICYQIKETIITCPHKPEGSFMVVKLNHLLEBDDIDDFCCKLAREESVILCFP
SVLGWMNVRTITACVPSSLQDLGRIKSFQORKNKSNSSDDC"
join(6518..6897,7029..7371,7479..7697,7784..7843,
8285..8440,8738..8831,9414..9547)
/gene="naat-A"
CDS
join(6518..6897,7029..7371,7479..7697,7784..7843,
8285..8440,8738..8831,9414..9547)
/gene="naat-A"
/codon_start=1
/product="nicotianamine aminotransferase"
/protein_id="BAAB7055.1"
/db_xref="GI:6469090"
/translation="MVHDSNGHGEAAHAAANGKSNHGHAHAAANGKSNHAAVENVNF
ARGKGLIATGAKNSRAIRYKISAEVESGPRVPLAHGDPSPVFAPTAVEAED
AAVALRTGQFNCPAGVGLPAARSAVAEHLSDGVPVYMLSADVDVLTAGTQALEVIL
PVLAOTAGANILLPSPGYPNVEARAAENREVRHFDLIPKGMHEIDDSLESADKNT
TAMVILNPNPFGSVYSDHLSKVAEVAEHLSDGVPVYMLSADVDVLTAGTQALEVIL
HAPVLSIGLSKSMIVFGWLGWVAVDPTKILEKTKISTINYLNVDTPATFVQ
EALPKILENTKADFFKRIIGLLKRESSEICYREIKENKVIITCPHKPEGSFMVVKLNHL
LLEEHDDINFCCKLAREESVILCFPSVLGMNVWRTITACVPSSLQDLGRIKSFQ
RNKKNKSNKSGC"
BASE COUNT 3082 a 2422 c 2366 g 3096 t
ORIGIN

Alignment Scores:
Pred. No.: 1.47e-131 Length: 10966
Score: 2484.00 Matches: 550
Percent Similarity: 51.30% Conservative: 1
Best Local Similarity: 51.21% Mismatches: 0
Query Match: 86.97% Indels: 523
DB: 8 Gaps: 6

US-10-019-783-2 (1-551) x AB024006 (1-10966)
QY 1 MetAlaThrValArgGlnSerAspGlyValAlaAlaAsnGlyLeuAlaValAlaAla 20
Db 650 ATGGCCACCGTACCGCAGACGAGTCCGCGGAAACGGCTTTCGCGTGGCCGAGCC 709
QY 21 AlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAlaValAsnGlyLysSerAsnGly 40
Db 710 GCGAACCGCAAGACGACCGCATCGCTGCGTCCGCGGTGAACGCGCAAGACGCGC 769
QY 41 HisGlyValAlaAlaAsnGlyLysSerAsnGlyHisGlyValAlaAlaAlaAla 60
Db 770 CATGCGTGGATGCGACGCGACGCGACGACGACGACGACGACGACGACGACGCG 829
QY 61 AsnGlyLysSerAsnGlyHisGlyValAlaAlaThrAlaAlaAlaAlaAlaAla 80
Db 830 AACGCGCAAGACGACCGCATCGCGGACCGCACTGCGGAAACGCGCCACGCGGAGCCACTCGC 889

```









QY 439 PheLysAlaIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLys 458  
 Db 1092 TTTCAGAAAGAACAGACTACTCAACATAATGTTGGTGTGATAGGCTCAAG 1151  
 QY 459 GluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLys 478  
 Db 1152 GACATCCCTGCTGCTGCTGCCAAGAACCTGAGTCTTCGCACTTACTTATTGACAAAG 1211  
 QY 479 LeuAsnLeuHisLeuLeuGluGluIleAspAspAspIleAspPheCysLysLeuAla 498  
 Db 1212 TTGGAGCTGCTATTGATGATGATATCAAGCAGCATAGATTTTTCGTAACACTGCC 1271  
 QY 499 LysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArg 518  
 Db 1272 AGAGAGAGAAATCTGCTGTTCTTACCAGGGATGCTCTGGGTTTCAAGAACTGGGTGAGG 1331  
 QY 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538  
 Db 1332 ATACCATCGGAGTCGAGCTCATATGCTTGAGATGCATTTGAGACACTGAAGGTTTC 1391  
 QY 539 CysGlnArgAsnLysLysArgAsnSerSerAsp 549  
 Db 1392 TGTACACGTCTATGCTCAAGAACAGACAGACAGAA 1424  
 RESULT 8  
 AF301900 1611 bp mRNA linear PLN 03-DEC-2000  
 LOCUS Arabidopsis thaliana Wassilewskija Ws-2 ROOTY/SUPERROOT1 (RTY/SUR1)  
 DEFINITION mRNA, complete cds.  
 ACCESSION AF301900  
 VERSION AF301900.1 GI:11527942  
 KEYWORDS Arabidopsis thaliana (chale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1 (bases 1 to 1611)  
 Gopalraj, M. and Olszewski, N.E.  
 AUTHORS The ROOTY/SUPERROOT1 gene of Arabidopsis encodes a putative  
 TITLE tyrosine aminotransferase  
 JOURNAL unpublished  
 REFERENCE 2 (bases 1 to 1611)  
 Gopalraj, M. and Olszewski, N.E.  
 AUTHORS Direct Submission  
 TITLE Submitted (02-SEP-2000) Plant Biology, University of Minnesota,  
 JOURNAL 1445 Gortner Ave, St. Paul, MN 55108, USA  
 FEATURES Location/Qualifiers  
 1..1611  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /note="ecotype: Wassilewskija Ws-2"  
 1..1611  
 /gene="RTY/SUR1"  
 56..1444  
 /gene="putative tyrosine aminotransferase"  
 /codon\_start=1  
 /product="ROOTY/SUPERROOT1"  
 /protein\_id="AAG37063.1"  
 /db\_xref="GI:11527943"  
 /translation="MSSEQPHANLAVPAFKTEKDPITQTRGSSVVRFGSDKA  
 STVLRGVIYMLPRVADVMNRDLPHKLTPEIDFLTAGCNOGIEIVFESLARPNANTLL  
 SYGPGAGILPARAVADVMNRDLPHKLTPEIDFLTAGCNOGIEIVFESLARPNANTLL  
 PRPGPHVDAAAYVSGLEVKFDLLPEKEWIEIDEGTEADEKTVAMVIVNPNPCG  
 NVVSHDLKVKVAFARLGMVLSDEVVDPTIEGDNPFVPMGFKAISIVPLLAGLSK  
 GWVVGKIGWIALNDREGVFETTKVLQSKQNLDTPTDPTATIIQALPAILEKADKN  
 FFAKKNLKHNDVLDVRUKDIPCVVCPKPKPSCTYLLTKLELSMDNKKDIDFCV  
 KLAEEENLFLPGDLGLKNWRRITIGVEAHMLEDALERLKGFCRTHAKKTETETESL  
 QALKLSNDNLEM"  
 154  
 /gene="RTY/SUR1"  
 variation

variation  
 185..235  
 /note="produces a stop codon which truncates the protein  
 at amino acid 32 in the alf1 allele"  
 /replace="a"  
 /gene="RTY/SUR1"  
 /note="deleted in the hls3 allele"  
 /replace="n"  
 BASE COUNT 466 a 352 c 374 g 419 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,378-55 Length: 1611  
 Score: 1124.50 Matches: 210  
 Percent Similarity: 65.39% Conservative: 98  
 Best Local Similarity: 44.59% Mismatches: 142  
 Query Match: 39.37% Indels: 21  
 DB: 4 Gaps: 4  
 US-10-019-783-2 (1-551) x AF301900 (1-1611)  
 QY 83 LysThrAsnGlyHisArg-----GluSerAsnGlyHisAlaGluAlaAlaAsp 98  
 Db 32 AAGAACTCAAGCACAGAGAGAGATGAGCGAAGAACCAACACACGCAATCTAGCGTT 91  
 QY 99 AlaAsnGlyGluSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGly 118  
 Db 92 CCCGGCTTAAACTGAGAAAGATCCCTAATACGCAACACGAAATGGTCAAGTAGC--- 148  
 QY 119 HisAlaAlaAlaAlaGluGluAlaValGluTrpAsnPheAlaGlyAlaLys 138  
 Db 149 -----GTTTGGCGTTTCGTTGGAAGTAT 172  
 QY 139 AspGlyValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrIle 158  
 Db 173 AAG-----GCACGAAAGCATCCACCGTAAGCTTAGAGGTGTCATCTATGCTC 223  
 QY 159 SerAlaSerValGlnGlyGlyProArgProValLeuProLeuAlaHisGlyAspPro 178  
 Db 224 TTCGCACTCGGCAAGACGTCATAGACCATTTTACCCCTCGGCACGTCACCT 283  
 QY 179 SerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaVal 198  
 Db 284 TCCGTCTACCTTGTCTCCGTACCTGTATCGAAGCTGAAGACGCGCTCGTCGAGCTCCT 343  
 QY 199 ArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAla 218  
 Db 344 CCTCCGCAAGGCAATCTTACGCTCCGAGCTGGGATTTCCCGCAGACGAGCC 403  
 QY 219 ValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspAspValPheLeu 238  
 Db 404 GTTGCTGATTATATGACCGAGATCTTCGACAAAGTTTACCGCTGAAGATATTTCTG 463  
 QY 239 ThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThrAlaGly 258  
 Db 464 ACCGCTGATGCAACCAAGGATAGATGCTGTTCGAATCGTTGGCTCGA---CCAAAC 520  
 QY 259 AlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsn 278  
 Db 521 GCAACATCTTCTCCACGCTCTGGCTTCCTCATACGACGCTCGCTGCTTACAGT 580  
 QY 279 ArgLeuGluValArgHisPheAspIleProAspLysGlyTyrGluIleAspIleAsp 298  
 Db 581 GGTCTCGAGGTTCGCAAGTTTGATCTTCTCCGCAAGAAAGATGGAGATGATCTTGA 640  
 QY 299 SerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsn 318  
 Db 641 GGTATCGAAGCATTCGACAGAGAAAACCTGGCTATGGTTGTTGTAATTAACCCCAAT 700  
 QY 319 ProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArg 338  
 Db 701 CCTCTGGAATGCTACTCTCAGACCATCTCAAAAGGTTGACAGACGCGTAGGAAG 760  
 QY 339 LeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaPro 358



```

Db 761 CTCGGATAATGGTATCTCAGACGAAGTATATGACCGAACTATATTTCGGAGACAATCCA 820
Qy 359 PheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSer 378
Db 821 TTGTTCCTCAATGGGAAGTTTCCTTCGATAGTCCTGTGATGACATGACAGGATATCT 880
Qy 379 LysSerTPIleValProGlyTIPArgLeuGlyTIPValAlaValTyAspProArgLys 398
Db 881 AAGGATGGTGTTCCTGGATGGAATTCGCTGATCGCTTGAATGATCCGAGGGC 940
Qy 399 IleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyLeuAsnValSerThrAsp 418
Db 941 GTTTCGAGACCAACCAAGGTGTACATCCATCAACAGAACTCTTTCAGCTAACTCCTGAC 1000
Qy 419 ProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhe 438
Db 1001 CTTGCCACATATATTCAGCTGCTCATCTCCAGGATCTCCGAGAACTCCGAGAACTTC 1060
Qy 439 PheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyLeuGlnIleLys 458
Db 1061 TTTCAGAAAGAGAACCAAGACTACTCAACATATATGTTGATTTGGTGTGATAGGCTCAAG 1120
Qy 459 GluAsnLysTyIleThrCysProHisLysProGluGlySerMetPheValMetValLys 478
Db 1121 GATATCCCTGTGTCTGTCTCCCAAGAACTGAGTCTTGCACTTACTTATTGACAAG 1180
Qy 479 LeuAsnLeuHisLeuLeuGluLeuGluLeuAspAspIleAspPheCysLysLeuAla 498
Db 1181 TTGGAGCTGTCACTGATGATATATCAAGGACGATATAGATTTTTCGTTAACTGGCC 1240
Qy 499 LysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArg 518
Db 1241 AGAGAGGAGATCTCGTGTCTTCTACCGGGGATGCTCTGGGTTTGAAGAACTGGATAGG 1300
Qy 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538
Db 1301 ATACCATCGGAGTGAAGCTCATATGCTTGAGATGACATTGACAGACTGAAGGTTTC 1360
Qy 539 CysGlnArgAsnLysLysArgAsnSerSerAsp 549
Db 1361 TGTAACGTCATCGCAAGAGAGACAGACAGAA 1393

RESULT 9
AF301898 1611 bp mRNA linear PIN 03-DEC-2000
LOCUS Arabidopsis thaliana Columbia ecotype RTY/SUR1 mRNA,
DEFINITION complete cds.
ACCESSION AF301898
VERSION AF301898.1 GI:11527938
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1611)
Gopalraj,M. and Olszewski,N.E.
The ROOTY/SUR1 gene of Arabidopsis encodes a putative
tyrosine aminotransferase
Unpublished
2 (bases 1 to 1611)
Gopalraj,M. and Olszewski,N.E.
Direct Submission
TITLE Submitted (02-SEP-2000) Plant Biology, University of Minnesota,
JOURNAL 1445 Gortner Ave, St. Paul, MN 55108, USA
LOCATION/Qualifiers
FEATURES
source
1..1611
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chrnosmes="II"
/map="between PHYB and CHL2"

```

```

/notes="ecotype: Columbia"
1..1611
/genes="RTY/SUR1"
56..1444
/genes="RTY/SUR1"
/notes="putative tyrosine aminotransferase"
/codon_start=1
/product="ROOTY/SUR1"
/protein_id="AAG37061.1"
/db_xref="GI:11527938"
/translation="MSEEPHANLAVPAKTEKEPITOTRNGOSSVRRGGSDKAAGA
STVTLGKVIYMLFDNCGKDVNKTLPLCHGDSVYPCPRCTCEAEADAVDVLRSKGN
SYGFGAGILPARRAVADYNRDLPHKLTEDILITAGCNGGIEIVFESLARPNAILL
PRGFFPHDARAVASGLVKRFDLLEPEKEWIDLEGEIAEADENTVAMVVPNNPCG
NVTSHDLKVAETAKLGMTWISDEVDTITFDGNPFVSMGFASIVPVLTLAGISK
FWVPGWKILGWALNDPEGVFTKVLQSKNDLVDTPDPATTIQAALPALEKADKN
GFAKKNKILKHNVDLVCDRLKIDPCVCPKPCSCYLLTKLELSMDNKIDIDFCV
KLARENLYPLGDLGLKXNWHIITIGVEAHMEDALERLKGFCFTRHAKKTETETESL
QALKSLDNMLEM"
399
/genes="RTY/SUR1"
/notes="Arg changed to Gln in sur1-4"
/replace="a"
692
/genes="RTY/SUR1"
/notes="Pro changed to Ser in rty-1"
/replace="t"
1253
/genes="RTY/SUR1"
/notes="Leu changed to Phe in sur1-3"
/replace="t"
1269
/genes="RTY/SUR1"
/notes="Gly changed to Gln in sur1-1"
/replace="a"
BASE COUNT 465 a 352 c 375 g 419 t
ORIGIN
Alignment Scores:
Pred. No.: 2.7e-55 Length: 1611
Score: 1123.50 Matches: 210
Percent Similarity: 65.39% Conservative: 98
Best Local Similarity: 44.59% Mismatches: 142
Query Match: 39.34% Indels: 21
DB: 8 Gaps: 4
US-10-019-783-2 (1-551) x AF301898 (1-1611)
Qy 83 LysThrAsnGlyHisArg-----GlusAsnGlyHisAlaGluAlaAlaAsp 98
Db 32 AAGAACTCAAAGCACAGAGAGAGATGAGCGAAGAACACACACGCAATCTAGCGTT 91
Qy 99 AlaAsnGlyGluSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGly 118
Db 92 CCGCGTTTAAACTGAGAAAGAGCCCATACCAACAGAAATGTCAGAGTAGC--- 148
Qy 119 HisAlaAlaAlaAlaAlaGluGluGluAlaValGluTrpAsnPheAlaGlyAlaLys 138
Db 149 -----GTTTGGCGTTTCGTTGGAGTGAT 172
Qy 139 AspGlyValLeuAlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyLysIle 159
Db 173 AAG-----GCAGCGAAAGCATCCACCGTACCGTTAGAGGTGTCATCTCATCTC 223
Qy 159 SerAlaSerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspPro 178
Db 224 TTCGACAACTGGCGCAAGAGAGCTCATAGACCATTTTACCTCCGCCACGGTAGCCT 283
Qy 179 SerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaVal 198
Db 284 TCCGCTACCTTCTGCTTCCGTACCTGTATCGAAGCTGAAGACCGCGTCGTCGACGCT 343
Qy 199 ArgThrGlyGlnPheAsnCysTyProAlaGlyValGlyLeuProAlaAlaAlaSerAla 218

```

```

Db 344 CGTCCGCAAGGCAATTTTACGGTCCGAGCTGGATCTCCCGCAAGCAGACC 403
Qy 219 ValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeu 238
Db 404 GTTGTGATTATGAACCGAGATCTTCGGCAAGTTTAAACGCTGAAGATATTTCTG 463
Qy 239 ThrAlaGlyGlyThrGlnAlaIleGluValIleLeuProValLeuAlaGlnThrAlaGly 258
Db 464 ACCGTGGATCAACCAAGGATGAGATCGTGTTCGAATCGTTCGCTCGA---CCAAAC 520
Qy 259 AlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsn 278
Db 521 GCACACATCTTCTCCAGCTCTGGTCTCTCTACAGGCTCGTCTGTACAGT 580
Qy 279 ArgLeuGluValArgHisPheAspLeuLeuProAspLysGlyTyrGluIleAspIleAsp 298
Db 581 GGTCTCGAGTTCGCAAGTTTCTCTTCCGAGAAAGATGGAGATTGATCTTGAA 640
Qy 299 SerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleLeuAsnProAsn 318
Db 641 GGTATCGAAGCATTCACAGAGAACACTGTGGCTATGTTGTTAATTAACCCCAACAT 700
Qy 319 ProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArg 338
Db 701 CCTCTGGAAATGTTCTACTCTCAGCACCATCTCAAAAAGTTTCAGAGACGCTAGGAAG 760
Qy 339 LeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaPro 358
Db 761 CTCGGGATAATGGTATCTCAGCAAGATATATGACCAACTATATTCGGAGACAATCCA 820
Qy 359 PheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSer 378
Db 821 TTGTGTTCAATGGGAAGTTTCTTCGATAGTCTCTGATTTGACACTAGCAGCATATCT 880
Qy 379 LysSerTrpIleValProGlyTyrPargLeuGlyTyrValAlaValTyrAspProArgLys 398
Db 881 AAGGATGGTGTGTTCTCGGATGGAATATGGCTGGATCGCTTGAATGATCCCGAGGGC 940
Qy 399 IleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAsp 418
Db 941 GTTTCGAGACCAACCAAGTGTACATCCATCAACAGAACTCTGAGAAAGCGCAAAACTTC 1060
Qy 419 ProAlaThrPheIleGlnAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhe 438
Db 1001 CCTGCCACAATAATTCAGGCTGCATCTCCAGCGATCTCGGAGAAAGCGCAAAACTTC 1060
Qy 439 PheLysAlaIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLys 458
Db 1061 TTGCAAGAGAACCAAGATCTCAACATATGTTGATTTGGTGTGTAGCTCAAG 1120
Qy 459 GluAsnLysTyrIleThrCysProHisLysProGlySerMetPheValMetValLys 478
Db 1121 GACATCCCTCGTGTCTGCTGCTCCCAAGAACTGAGTCTTGCATCTTACTTATTCACAAG 1180
Qy 479 LeuAsnLeuHisLeuLeuGluGluIleAspAspIleAspPheCysLysLeuAla 498
Db 1181 TTGGAGCTGTCTATGATGATAATATCAAGACCATATAGATTTTTCGTAACATCGCC 1240
Qy 499 LysGluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyrValArg 518
Db 1241 AGAGAGGAGATCTCGTGTCTTACAGGGATGCTCTGGGTTTGAAGACTGGATGAGG 1300
Qy 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538
Db 1301 ATAACCATCGGAGTCGAAGCTTCATATGCTTGAGGATGCATCTTGAGACTCAAGGGTTC 1360
Qy 539 CysGlnArgAsnLysLysArgAsnSerSerAsp 549
Db 1361 TGTACAGTCTATGCCAAGAACACAGACAGAA 1393

```

RESULT 10  
AF301899

```

LOCUS AF301899 1611 bp mRNA linear PLN 03-DEC-2000
DEFINITION Arabidopsis thaliana Landsberg erecta ROOTY/SUPERROOT1 (RTY/SUR1)
mRNA, complete cds.
ACCESSION AF301899
VERSION AF301899.1 GI:11527940
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eucotsids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1. (bases 1 to 1611)
Gopalraj, M. and Olszewski, N.E.
The ROOTY/SUPERROOT1 gene of Arabidopsis encodes a putative
tyrosine aminotransferase
Unpublished
REFERENCE 2. (bases 1 to 1611)
Gopalraj, M. and Olszewski, N.E.
Direct Submission
TITLE Arabidopsis thaliana
JOURNAL Submitted (02-SEP-2000) Plant Biology, University of Minnesota,
1445 Gortner Ave, St. Paul, MN 55108, USA
FEATURES
source
1. ..1611
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/notes="ecotype: Landsberg erecta"
gene
1. ..1611
/gene="RTY/SUR1"
CDS
56. ..1444
/gene="RTY/SUR1"
/notes="putative tyrosine aminotransferase"
/codon_start=1
/product="ROOTY/SUPERROOT1"
/protein_id="AAG37062.1"
/db_xref="GI:11527941"
/translation="MSEEPHANLAVPAFKTEKDPVDTQNGQSVWFGSGDKAARKA
SVTVLRGVLVFLFDCGDKVNTIPLGHGDPSPVPCFCTCIEAEDAVVDVLRSGKN
SGPCAGILPARRAVADYMRDLPHKLTPEIDFILTACNGQIEIVFESLAPNANILL
PRPGSDARAAVSGLEVRKFDLLPEKEWIDLEGIEAIDEKTVAMVNPNNPCK
NYSHDHUKVAETARKLIMVISEVDYRTIFGDNPEVPMKGFASIVPVITLAGISK
GWVPGWIKHNDLVCDLKDIPCVCQKPKPESCTYLLTLELSLMDNIKDDIDFCV
FFAKNKLKHNDLVCDLKDIPCVCQKPKPESCTYLLTLELSLMDNIKDDIDFCV
KLARENILVFLPGVALGLKRWIRITIGVEAHMLEDALERLKGFCFTRHAKKTETESL
QALKLSDNNLEM"
variation
1294
/gene="RTY/SUR1"
/notes="produces a stop codon which truncates the protein
at amino acid 412 in rty-4"
/replace="a"
BASE COUNT 464 a 351 c 375 g 421 t
ORIGIN
Alignment Scores:
Pred. No.: 2,7e-55 Length: 1611
Score: 1123.50 Matches: 210
Percent Similarity: 65.39% Conservative: 98
Best Local Similarity: 44.59% Mismatches: 142
Query Match: 39.34% Indels: 21
DB: 4 Gaps: 4
US-10-019-783-2 (1-551) x AF301899 (1-1611)
Qy 83 LysThrAsnGlyHisArg-----GlusAsnGlyHisAlaGluAlaAlaAsp 98
Db 32 AAGAACTCAAGACACAGAGAGAGATGAGCAAGAACACACACCGCAATCTAGCGTT 91
Qy 99 AlaAsnGlyGluSerAsnGluHisAlaGluAspSerAlaAlaAsnGlyGluSerAsnGly 118
Db 92 CCGCGGTTTAAACATGAGAAAGATCCCGTAACGCAACGCAAAATGGTCAAGTAGC--- 148
Qy 119 HisAlaAlaAlaAlaGluGluGluAlaValGluTyrAsnPheAlaGlyAlaLys 138

```

Db 149 -----GTTTGGCGTTTCGGTGGAGTGAT 172  
 QY 139 AspGlyValLeuAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrIle 158  
 Db 173 AG-----GCAGCAAGACATCCACCGTACGCTTAGAGTGTATCTATCATGCTC 223  
 QY 159 SerAlaSerValGlnGlyGlyProArgProValLeuProLeuAlaHisGlyAspPro 178  
 Db 224 TTCGCACAACTGGCGCAAGACGCTCAATAAGACCACTTTTACCCCTCGGCACGGTGACCCCT 283  
 QY 179 SerValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaVal 198  
 Db 284 TCCGCTACCTTGTCTCGTACCTGTATCGAAGCTGAAGACGCGCTGTGCGACGCTCCTT 343  
 QY 199 ArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaArgSerAla 218  
 Db 344 CGCTCCGCAAGCAAGCAATCTTACGCTCCCGAGCTGGGATCTCCCGCAGACGAGCC 403  
 QY 219 ValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeu 238  
 Db 404 GTGTGCTGATTATATGAACCGAGATCTCCGCACAAAGTTAACGCTGAAGATATTTTCTG 463  
 QY 239 ThrAlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGly 258  
 Db 464 ACCGCTGGATGCAACCAAGGATAGATCGTGTTCGATCGTGGCTCGA---CCAAAC 520  
 QY 259 AlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAlaPheAsn 278  
 Db 521 GCAACATCTTGTCTCCCGCTCTCGCTTCCCTCATACGACGCTCGTGTCTTACAGT 580  
 QY 279 ArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluAlaAspIleAsp 298  
 Db 581 GGTCTCGAGGTTCCGAAGTTTGATCTTCTCCGAGAAAGATGGAGATTTGATCTTGA 640  
 QY 299 SerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsn 318  
 Db 641 GGTATCGAAGCATTGCCAGACGAGAAAATCTGTGGCTATGGTTGTAATTAAACCCCAACAAT 700  
 QY 319 ProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArg 338  
 Db 701 CCCTGTGAAATGTCTCTCACACCATCTCAAAAGGTTGCAGACGCGTAGGAAG 760  
 QY 339 LeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaPro 358  
 Db 761 CTCGGATAATGGTGATCTCAGACGAGATATATGACCACTATATTCGGAGACAATCCA 820  
 QY 359 PheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSer 378  
 Db 821 TTTGTCCAAATGGGGAAGTTTGTCTCGATAGTCCCTGTATTGACACTAGCAGGATATCT 880  
 QY 379 LysSerTrpIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLys 398  
 Db 881 AAGGATGGTGTTCCTGGATGGAAATTTGCTGGATCGCCTTGAATGATCCCGAGGCG 940  
 QY 399 IleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAsp 418  
 Db 941 GTTTTCGAGACCAACCAAGGTGTACATCCATCAACAGAGATCTTCAGTAACCTCTGAC 1000  
 QY 419 ProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhe 438  
 Db 1001 CCTGCCACAAATATTCAGCGCTGCACTCCACGCACTCTCCGAGAAAGCGGACAAATCTC 1060  
 QY 439 PheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLys 458  
 Db 1061 TTTGCAGAGAGAACCAAGATCTCAACATATGTTGATTTGGTGTGTGATAGCTCAAG 1120  
 QY 459 GluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLys 478  
 Db 1121 GACATCCCTGTGCTGTCTGCCAAGAAACCTGAGCTTTCACCTTACTTATTCACAAAG 1180  
 QY 479 LeuAsnLeuHisLeuLeuGluIleAlaAspAspIleAspPheCysLysLeuAla 498  
 Db 1181 TTGGAGCTGTATTGATGGATAATATCAAGACGATATAGATTTTTCGTTAAAGCTGGCC 1240

QY 499 LysGluIleSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrpValArg 518  
 Db 1241 AGAGAGAGATCTCTGGTTTTCACAGGATGCTCTGGTTTGAAGAACTGATGAGG 1300  
 QY 519 IleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPhe 538  
 Db 1301 ATAACCATCGGATCGAAGCTCATATGCTTGAAGATGCATCTTGAGAGACTGAAGGGTTTC 1360  
 QY 539 CysGlnArgAsnLysLysArgAsnSerSerAsp 549  
 Db 1361 TGTACAGTCATCGCAAGAAGACAGACAGAA 1393

## RESULT 11

LOCUS AY050987 1683 bp mRNA linear PLN 18-SEP-2002  
 DEFINITION Arabidopsis thaliana putative tyrosine aminotransferase (At2g20610)  
 mRNA, complete cds.

ACCSSION AY050987

VERSION AY050987.1 GI:15293106

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

## REFERENCE

1 (bases 1 to 1683)  
 Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
 Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
 Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
 Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
 Davis, R.W., Ecker, J.R. and Theologis, A.  
 Arabidopsis Full Length cDNA Clones

## TITLE

Unpublished

## JOURNAL

2 (bases 1 to 1683)

## REFERENCE

Yamada, K., Liu, S.X., Pham, P.K., Banh, J., Banno, P., Dale, J.M.,  
 Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L.,  
 Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C.,  
 Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M.,  
 Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
 Southwick, A., Tracy, S.E., Shinozaki, K., Davis, R.W., Ecker, J.R. and  
 Theologis, A.

## TITLE

Direct Submission

## JOURNAL

Submitted (07-AUG-2001) Plant Gene Expression Center, 800 Buchanan  
 Street, Albany, CA 94710, USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN  
 Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
 Hayashizaki, Y. and Shinozaki, K.

## COMMENT

The Salk, Stanford, PGEN (SSP) Consortium members carried out the  
 sequencing and annotation of the RAPL cDNAs: Yamada, K., Liu, S.X.,  
 Pham, P.K., Banh, J., Banno, P., Dale, J.M., Goldsmith, A.D.,  
 Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
 Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H.,  
 Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B.,  
 Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P.,  
 Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Yamada, K. (SSP/PGEN) and Seki, M. (RIKEN GSC) contributed equally to  
 this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEN)  
 contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
 genome submitted to GenBank.

## FEATURES

source

1. .1683

/organism="Arabidopsis thaliana"

```

/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAF107-08-L08 (R10872)"
/note="This clone is in a modified pBluescript vector
(FLC-1) as a BamHI/XhoI insert."
ecotype: Columbia"
1..1683
/gene="At2g20610"
5'UTR
1..87
/gene="At2g20610"
misc_difference 2
/gene="At2g20610"
/note="not present in genomic sequence"
misc_difference 21..22
/gene="At2g20610"
/note="compared to genomic sequence"
/replace="ttc"
88..1476
/gene="At2g20610"
/codon_start=1
evidence=experimental
/product="putative tyrosine aminotransferase"
/protein_id="AAK93664.1"
/db_xref="GI:15293107"
/translation="MSEEPHANLAVPAFKTEKEPIQTNRGOSVWRFSGSKAARA
STVLRGVLYVLFNCCGKDVNKTILPLGHGDPSPVYPCFRTICIEAEDAVDVLRSKGN
SYGCGAGILPARRAVAYMNRDLPHKLTLPEDIFLTAGNCGIETVPESLARPNAILL
PRGPFYDARAASGLVEKFDLPPEKEWIDLEGEALADENTVAMVINPNPCG
NVYSHDKVAETARKLGINVIDEYDRTIFGDNFVSGKFPASIVPVLTAGISK
GWVPGWKIHALNDPEGEFTTKVQSKQNLDVDPDPATIIQAALPAILEKADKN
FFAKNKILHNLVDLVDRLKDI PCVVCPPKPSCTVLLKLSLSMDNKKDIDPCV
KLARENILVFPGGALGKWNWRITIGVEAHMLEDALERLKGFCFTRHAKKTETESL
QALKSLDNMLEM"
1477..1683
3'UTR
/gene="At2g20610"
misc_difference 1667
/gene="At2g20610"
/note="not present in genomic sequence"
BASE COUNT 497 a 356 c 384 g 446 t
ORIGIN

Alignment Scores:
Pred. No.: 2,856-55 Length: 1683
Score: 1123.50 Matches: 210
Percent Similarity: 65.39% Conservative: 98
Best Local Similarity: 44.59% Mismatches: 142
Query Match: 39.34% Indels: 21
DB: 8 Gaps: 4

US-10-019-783-2 (1-551) x AY050987 (1-1683)
QY 83 LysThrAsnGlyHisArg-----GluSerAsnGlyHisAlaGluAlaAasp 98
DB 64 AAGAACTCAAGCAGCAGAGAGAGATGAGCGAAGACACACACACCAATCTAGCGGTT 123
QY 99 AlaAsnGlyGluSerAsnGluHisAlaGluAaspSerAlaAAsnGlyGluSerAsnGly 118
DB 124 CCGCGGTTTAAACTGAGAAAGAGAGCCATACCGCAACACCAATGTCAGAGTAGC 180
QY 119 HisAlaAlaAlaAlaGluGluGluAlaValGluTTPAsnPheAlaGlyAlaLys 138
DB 181 -----GTTGGCGTTTCGGTGGAGTAGT 204
QY 139 AspGlyValLeuAlaAlaThrGlyAlaAsnMetSerleArgAlaAlaAArgTyrLysile 158
DB 205 AAG-----GCAGCGAAAGCATCCACCGCTTAACGGTTAGAGGTGTCATCATGCTC 255
QY 159 SerAlaSerValGlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAaspPro 178
DB 256 TTCGCAACTCGGCGAAAGAGCTCAATNAGACCAATTTTACCCCTCGGCCACGGTGACCT 315
QY 179 SerValPheProAlaPheArgThrAlaValGluAlaGluAaspAlaValAlaAlaVal 198

```

```

DB 316 TCGGCTACCCCTGCTTCGGTACCTGTATCGAAGCTGAAGCGCGTCTGTCGACGCTCTT 375
QY 199 ArgThrGlyGlnPheAsnCysTyrProAlaGlyValGlyLeuProAlaAlaAArgSerAla 218
DB 376 CGCTCCGGCAAGGCAATTTTACGGTCCCGAGCTGGGATTTCCCGCAGAGAGGCC 435
QY 219 ValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAaspValPheLeu 238
DB 436 GTTGCTGATTATGACCGAGATCTTCGGCAAGTTTAAACGCTCGAAGATATTTCTG 495
QY 239 ThrAlaGlyGlyThrGluAlaAlaGluVallelleProValleuAlaAlaGluThrAlaGly 258
DB 496 ACGCTGGATGCAACCAAGGATAGATCGTGTTCGAATCGTGGCTCGA---CCAAAC 552
QY 259 AlaAsnLeuLeuProArgProGlyTyrProAsnTyrGluAlaAArgAlaAlaPheAsn 278
DB 553 GCAACATCTTCTCCCGCTCTGCTCTCCCTACTACGAGCTCGTCTGCTTACAGT 612
QY 279 ArgLeuGluValArgHisPheAspLeuLeuProAspLysGlyTyrGluLeuAaspLeu 298
DB 613 GGTCTCGAGTTCCGCAAGTTTGATTTCTTCCGAGAAAGATGGGAGATTGATCTTGA 672
QY 299 SerLeuGluSerIleAlaAaspLysAsnThrThrAlaMetVallelleAsnProAsnAsn 318
DB 673 GGTATCGAAGCATTGACAGCAGAACACTGTGGCTATGTTGTTAATTAACCCCAACAAT 732
QY 319 ProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArg 338
DB 733 CCTGTGGAAATGTCTACTCTCACCACTCTCAAAAAGTTGACAGACCGCTAGGAAG 792
QY 339 LeuGlyLeuVallelleAlaAaspGluValTyrGlyLysLeuValleuLysSerAlaPro 358
DB 793 CTCGGATAATGTTGATCTCAGACGAAAGTATATGACCGAATATATTCGGAGACAATCCA 852
QY 359 PheIleProMetGlyValPheGlyHisIleThrProValleuSerIleGlySerLeuSer 378
DB 853 TTTGTTCAATGGGAAGTTTCTTCGATAGTCCCTGTATTGACACTAGCAGGATATCT 912
QY 379 LysSerTyrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLys 398
DB 913 AAGGATGGGTTGTTCTCGATGGAATGCTGGATCGCTTGAATGATCCCGAGGCG 972
QY 399 IleLeuGlnGluThrLysIleSerThrSerleThrAsnTyrLeuAsnValSerThrAsp 418
DB 973 GTTTTCGAGACCAACCAAGGTGTACATCCATCAACACAAGAATCTTGACGTAACTCCTGAC 1032
QY 419 ProAlaThrPheIleGluAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAaspPhe 438
DB 1033 CTGCGCACAAATATTCAGCTGCACTTCCAGGATCTTCAGGATCTTCGGAGAACCGGACAAACCTTC 1092
QY 439 PheLysAlaIleIleGlyLeuLysGluSerSerGluIleCysTyrLysGlnIleLys 458
DB 1093 TTTTCAAAAGAACACACAGATACTCAACATAATGTTGATTGTTGGTGTGTATAGGCTCAAG 1152
QY 459 GluAsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLys 478
DB 1153 GACATCCCTGCTGCTGCTGCTCCCAAGAACCTTGAGTCTTGCATCTTACTTATTTGACAAG 1212
QY 479 LeuAsnLeuHisLeuLeuGluIleAaspAspIleAaspPheCysCysLysLeuAla 498
DB 1213 TTGAGCTGCTCATTTGATTAATATCAAGGACGATATAGATTTTTCGTAATACTGGCC 1272
QY 499 LysGluGluSerVallelleuCysProGlySerValleuGlyMetAlaAsnTrpValArg 518
DB 1273 AGAGAGGAGAATCTCGTGTCTTACCGGGGATGCTCTGGGTTTGAAGAATCGATGAGG 1332
QY 519 IleThrPheAlaCysValProSerSerLeuGluAspGlyLeuGlyArgIleLysSerPhe 538
DB 1333 ATAACCATCGGAGTCGAAGCTCATATGTTGAGGATGCACCTTGAGACTGAAGGTTTC 1392
QY 539 CysGlnArgAsnLysLysArgAsnSerSerAsp 549

```

```

Db      1393 TGTACAGTCATGCCAAGACAGACAGACAGAA 1425
RESULT 12
AX653224
LOCUS      920 bp      DNA      linear      PAT 22-MAR-2003
DEFINITION      Sequence 3094 from Patent WO03000898.
ACCESSION      AX653224
VERSION      AX653224.1 GI:29156038
KEYWORDS
SOURCE      Oryza sativa
ORGANISM      Oryza sativa
REFERENCE
AUTHORS      Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
              Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE      Plant genes involved in defense against pathogens
JOURNAL      Patent: WO 03000898-A 3094 03-JAN-2003;
              Syngenta Participations AG (CH)
FEATURES
            source
            1..920
            /organism="Oryza sativa"
            /mol_type="genomic DNA"
            /db_xref="taxon:4530"
BASE COUNT      275 a 184 c 205 g 256 t
ORIGIN
Alignment Scores:
Pred. No.:      1,75e-55      Length:      920
Score:          1121.50      Matches:      220
Percent Similarity: 75.15%      Conservative: 28
Best Local Similarity: 66.67%      Mismatches: 31
Query Match:      39.27%      Indels: 51
DB:              6      Gaps: 2
US-10-019-783-2 (1-551) x AX653224 (1-920)
QY      217 SerAlaValAlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspVal 236
Db      67 AGTGTGTGGCAGATCAATTTGTCACGCACCTCCCATACAGCTATCTCTGATGACATC 126
QY      237 PheLeuThrAlaGlyGlyThrGlnAlaIleGluValIleIleProValLeuAlaGlnThr 256
Db      127 TTCTAACCGCTGGAGGAACTCAGCCATCAGGTCGTAATCTCAATCTTGCCTCA 183
QY      257 AlaGlyAlaAsnIleLeuLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaAla 276
Db      184 CTGGCACAACATATGCTTCTAGACCGTACCCAACTATGAAGCTCGAGCGCG 243
QY      277 PheAsnArgLeuGluValArgHisPheAspLeuIleProAspLysGlyTrpGluIleAsp 296
Db      244 TTCAACAACCTTGAA----- 258
QY      297 IleAspSerLeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnPro 316
Db      258 ----- 258
QY      317 AsnAsnProCysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAla 336
Db      259 -----GTGGCAGAGGTAGCA 273
QY      337 LysArgLeuGlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySer 356
Db      274 AGGAAGCTTGGATATTTGGTAATTACTGATGAGGTGATGTGTAATTTGGTTTGGGAGT 333
QY      357 AlaProPheIleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySer 376
Db      334 TCCCATTTGTCCTCAATGGGTGCTTTGGGCACATCGTACCATAATTAAACCATAGGATCG 393
QY      377 LeuSerLysSerTrpIleValProGlyTrpArgLeuGlyTrpValAlaValTyrAspPro 396
Db      394 CTATCAAGAGGTGGATAGTCCGGGATGGCGACTTGGTTGGGTAGCAATATGTGACCCC 453
QY      397 ArgLysIleLeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSer 416
Db      454 AAGAAGACTCTACAAGAAACCAAGATTGCAACATAATTACTAATTTCTTAAATGTTCA 513
QY      417 ThrAspProAlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGlu 436
Db      514 ACTGATCCAGCAACTTTCATTTCAGGAGCTCTACCGAATATTCTTAAGAATACCAAGAA 573
QY      437 AspPhePheLysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGln 456
Db      574 GAATTTCTTAAAGAGGATAATTGTTTGTTCACGAAACATCAGATATTTGCTATAGAGGA 633
QY      457 IleLysGluAsnLysTyrIleThrCysProHisLysProGluLysProGluLysSerMetPheValMet 476
Db      634 ATAAAGATATTAAATGCAATCATCTTCTCTCAACGCCGAGGATCCATGTTTGTGATG 693
QY      477 ValLysLeuAsnLeuHisLeuLeuGluIleAspAspAspIleAspPheCysCysLys 496
Db      694 GTGAAATTAACCTATATCTTTTGGAGGAATCCATGATGATGTTGATTTTGTGCCAA 753
QY      497 LeuAlaLysGluLysSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTrp 516
Db      754 CTTCGAAAGAGAGTCGGTGATCTTTGCCAGGAGGTGCTGGGATGAAGAATTGG 813
QY      517 ValArgIleThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLys 536
Db      814 GTTCGCATTACTTTTGTATTGATTCACTTCTCTCTGATGTTCTTGAGAGGATCAAA 873
QY      537 SerPheCysGlnArgAsnLysLysArgAsn 546
Db      874 TCCTTCTGCCAAGGACACAGAGAAAGAAAAC 903
RESULT 13
AX575731
LOCUS      1389 bp      DNA      linear      PAT 07-JAN-2003
DEFINITION      Sequence 9 from Patent WO02072848.
ACCESSION      AX575731
VERSION      AX575731.1 GI:27552220
KEYWORDS
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM      Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS      Badur, R. and Geiger, M.
TITLE      Increase in the vitamin E content in organisms due to an increase
              in the tyrosine aminotransferase activity
JOURNAL      Patent: WO 02072848-A 9 19-SEP-2002;
              Sungene GmbH & Co. KGAA (DE)
FEATURES
            source
            1..1389
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /db_xref="taxon:3702"
            /notes="unnamed protein product"
            /codon_start=1
            /protein_id="CAD60072.1"
            /db_xref="GI:27552221"
            /translation="MSEEOCHANLAVPAFKTEKDPVTQTQNGQSSVVRFGGDKAAKA
              SVTVRGVLYMLFDNCSKVNTKILPLGHGDSVYPCFTCIERADAVYDVLRSKGN
              SVQPGAGILPARRAVADYNNRDLPHKLPEDIPLTAGCNGGLEIVFESLARNANILL
              NVYSHDHLKKAETAGVLEVRKFDLPEKEWIDLEGEIADEKTVAMVYVLPNPGG
              NVYSPGWKTIQIALNDPEGVFTKVLQSIKQNLDPDPTATIIQAALPAILEKADRN
              FPAKNIKLKHNVDLVCDLKDIPCVCVCFKPSCTYLLTLELSLMDNKKDIDPCV
              KLARENVLVPLGDLGLKQNWTRITIGVEAHMLDALERLKGFCFTHAKKTETESL
              QALKSLDNKLEM"
BASE COUNT      391 a 322 c 336 g 340 t
ORIGIN
```

## Alignment Scores:

Pred. No.: 2,91e-55 Length: 1389  
 Score: 1121.50 Matches: 207  
 Percent Similarity: 65.73% Conservative: 96  
 Best Local Similarity: 44.90% Mismatches: 141  
 Query Match: 39.27% Indels: 17  
 DB: 6 Gaps: 3

US-10-019-783-2 (1-551) x AX575731 (1-1389)

QY 89 GluSerAsnGlyHisAlaGluAlaAlaAspAlaAsnGlyGluSerAsnGluHisAlaGlu 108  
 DB 7 GAAGAACAAACACACGCAATCTAGCGTTCCCGGGTTTAAACTGAGAAAGATCCCGTA 66  
 QY 109 AspSerAlaAlaAsnGlyGluSerAsnGlyHisAlaAlaAlaAlaGluGluGlu 128  
 DB 67 AGCAAAACGCAAAATGGTCAAGTAGC----- 93  
 QY 129 AlaValGluTyrAsnPheAlaGlyAlaLysAspGlyValLeuAlaAlaThrGlyAlaAsn 148  
 DB 94 -----GTTTCGGGTTTCGGTGAAGTAGC-----GCAGCAAGATCCACC 138  
 QY 149 MetSerIleArgAlaIleArgTyrLysIleSerAlaSerValGlnGluLysGlyProArg 168  
 DB 139 GTAAAGCTTAGAGGTGTCATCTACATGCTCTTCGACAACTCGACAAAGAGCGTCAATAAG 198  
 QY 169 ProValLeuProLeuAlaHisGlyAspProSerValPheProAlaPheArgThrAlaVal 188  
 DB 139 ACCATTTTACCCTCGGCACGGTGACCTTCGGTCTACCTTCCTCCGTACCTGTATC 258  
 QY 189 GluAlaGluAspAlaValAlaAlaAlaValArgThrGlyGlnPheAsnCysTyrProAla 208  
 DB 259 GAAAGCTGAAGACCGCGTCTCGACGCTCTCGCTCCGCAAGCAATCTTACGGTCCC 318  
 QY 209 GlyValGlyLeuProAlaAlaArgSerAlaValAlaGluHisLeuSerGlnGlyValPro 228  
 DB 319 GGAGCTGGATCTCTCCCGCAAGACGAGCGCTGCTGATATATGAACCGAGATCTCCG 378  
 QY 229 TyrMetLeuSerAlaAspValPheLeuThrAlaGlyGlyThrGlnAlaIleGluVal 248  
 DB 379 CACAAAGTTAACCGCGAAGATATTTTCTGACCGCTGGATCAACCAAGGATAGATC 438  
 QY 249 IleIleProValLeuAlaGlnThrAlaGlyAlaAsnIleLeuProArgProGlyTyr 268  
 DB 439 GTGTTTGAATCGTTGGCTCGA---CCAAACGCAACATCTTGTCTCCACGCTCTGGCTTC 495  
 QY 269 ProAsnTyrGluAlaArgAlaAlaPheAsnArgLeuGluValArgHisPheAspLeuIle 288  
 DB 496 CCTCATTAAGACGCTCGTCTGTACAGTGGTCTCGAGGTTCGCAAGTTGATCTTCT 555  
 QY 289 ProAspLysGlyTyrGluAlaAspIleAspSerLeuGluSerIleAlaAspLysAsnThr 308  
 DB 556 CCCGAAAGAAATGGAGATTCATCTTGAAGGTATCGAAGCCATTCAGACGAGAAAAC 615  
 QY 309 ThrAlaMetValIleLeuAsnProAsnProCysGlySerValTyrSerTyrAspHis 328  
 DB 616 GTGGCTATGTTGTAATTAACCCCAACATCCCTGTGGAATGTCTACTCTCAGACCAT 675  
 QY 329 LeuSerLysValAlaGluValAlaLysArgLeuGlyIleLeuValIleAlaAspGluVal 348  
 DB 676 CTCAAAAGGTTCAGACACGCTAGGAAGTCGGGATATGTTGATCTCAGACCAAGTA 735  
 QY 349 TyrGlyLysLeuValLeuGlySerAlaProPheIleProMetGlyValPheGlyHisIle 368  
 DB 736 TATGACCAATATATTCGGAGACAATCCATTTGTTCAATGGGAAGTTTGTCTTCGATA 795  
 QY 369 ThrProValLeuSerIleGlySerLeuSerLysSerTrpIleValProGlyTyrArgLeu 388  
 DB 796 GTCCCTGTATTGACATAGCAGGATATCTAAGGATGGGTGTTCTCTGGATGAAAAAT 855  
 QY 389 GlyTrpValAlaValTyrAspProArgLysIleLeuGlnGluThrLysIleSerThrSer 408  
 DB 856 GGCTGGATTCCTTGATTCCTCGAGGGCGGCTTTTCGAGACCCCAAGGTGTACAAATCC 915

QY 409 IleThrAsnTyrLeuAsnValSerThrAspProAlaThrPheIleGlnAlaAlaLeuPro 428  
 DB 916 ATCAACACAGAAATCTTGACGTAACTCTGACCTGCCACAAATAATTCAGGCTGCATTTCCA 975  
 QY 429 GlnIleLeuGluAsnThrLysGluAspPhePheLysAlaIleIleGlyLeuLeuLysGlu 448  
 DB 976 GCGATCTCGAGAAACGCGACAAACATCTTTGCAAGAGAAACAAGATCTCAACAT 1035  
 QY 449 SerSerGluIleCysTyrLysGlnIleLysGluAsnLysTyrIleThrCysProHisLys 468  
 DB 1036 AATGTTGATTTGGTGTGTAGTCAAGGATCCCTGCTGTCTGTCTGCTCCCAAGAAA 1095  
 QY 469 ProGluGlySerMetPheValMetValLysLeuAsnLeuHisLeuGluGluIleAsp 488  
 DB 1096 CTTGAGTCTTGCACTTACTTATTGCAAGTTGGAGCTGTCTACTGATGGATAATATCAAG 1155  
 QY 489 AspAspIleAspPheCysCysLysLeuAlaLysGluSerValIleLeuCysProGly 508  
 DB 1156 GACGATATAGATTTTTCGTAATACTGCCAGAGAGGAAATCTGCTTCTACAGGG 1215  
 QY 509 SerValLeuGlyMetAlaAsnTrpValArgIleThrPheAlaCysValProSerSerLeu 528  
 DB 1216 GATGCTCTGGTGTGAAGAACTGGACGAGGATTAACCTCGGAGTGAAGCTCATATGCTT 1275  
 QY 529 GlnAspGlyLeuGlyArgIleLysSerPheCysGlnArgAsnLysLysArgAsnSerSer 548  
 DB 1276 GAGGATGCATTCAGAGACTGAAGGTTTCTGTACACGTCATGCCAAGAGACAGACACA 1335  
 QY 549 Asp 549  
 DB 1336 GAA 1338

RESULT 14  
 AX575733 1243 bp DNA linear PAT 07-JAN-2003  
 LOCUS Sequence 11 from Patent WO02072848.  
 DEFINITION AX575733  
 ACCESSION AX575733.1 GI:27552222  
 VERSION AX575733.1  
 KEYWORDS Arabidopsis thaliana (thale cress)  
 SOURCE Arabidopsis thaliana  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE  
 1 Badur, R. and Geiger, M.  
 AUTHORS Increase in the vitamin E content in organisms due to an increase  
 in the tyrosine aminotransferase activity  
 TITLE Patent: WO 02072848-A 11 19-SEP-2002;  
 JOURNAL Sungene GmbH & Co. KGAA (DE)  
 FEATURES Location/Qualifiers  
 source 1..1243  
 /organism="Arabidopsis thaliana"  
 /mol\_type="Genomic DNA"  
 /db\_xref="taxon:3702"  
 1..>1242  
 /note="unnamed protein product"

CDS  
 /codon\_start=1  
 /protein\_id="CAD60073.1"  
 /db\_xref="GI:27552223"  
 /translation="MENGATTTSTTIKGLISLLMESITTEDEGGKRVISLGMGDP  
 LVSCEPTQVSLQAVSDLSLNKPHGVSPTVGLPOARAJAEVLSRDLVYKLSDDDF  
 IYSGCTQADVALDSMLAPRANILLPRGFIYELCAKFRHLYVYVDLLENGWEID  
 LDVALADENTVALVWNPNGNVSQYHLMKIAESAKKGLFLVIADSVYGLAF  
 GSKPVPVFGVSIVPVLTLGSLKRWIVPGRLGWFTVTPDPSFKDKPIERFKKY  
 FTLGGPATFIQAAPVTLLEQTSDFPKKTLNSLNSIDICCDWKKEIPCDSHRPE  
 SEAMWVNLISLDDVSDIDFCFKLARESVILLPGTAVGLKNLWLRITFAADATSI  
 HEAFKRIKCFVLRHAKTQPTI"

BASE COUNT 323 a 258 c 291 g 371 t  
 ORIGIN  
 Alignment Scores:



LDAVEALADENTVALVWPGNVCNIVSYQHLMKIAESAKKLGLFVLVIADEVYGHAF  
GSKPVPNGVFGSTVPVLTGLSLKRWIVPGWRLGVFTTDPGSGFKDPKIIERFKKY  
FDILGGPATFTCAAVPTILEOTDESFFKKTLSLNKSSDICCDDKEIPCIDSSHPPE  
GSMAMWKLNLISILEDVSDDDDFEKLAREBSVILLPGTAVGLKNWLRITFAADATSI  
BEAFRIKCFYLRHAKQYPTI"  
1246..1276  
/gene="At5g53970"

3'UTR

BASE COUNT 331 a 259 c 299 g 387 t

ORIGIN

## Alignment Scores:

Pred. No.:	3.64e-55	Length:	1276
Score:	1119.00	Matches:	218
Percent Similarity:	71.11%	Conservative:	70
Best Local Similarity:	53.83%	Mismatches:	113
Query Match:	39.18%	Indels:	4
DB:	8	Gaps:	2

US-10-019-783-2 (1-551) x BT001912 (1-1276)

QY	143	AlaAlaThrGlyAlaAsnMetSerIleArgAlaIleArgTyrLysIleSerAlaSerVal	162
DB	13	GCAACGACGACGACACATACCATCAAGGGATTCTGAGTTTCTAATCGAAAGCATC	72
QY	163	-----GlnGluLysGlyProArgProValLeuProLeuAlaHisGlyAspProSer	179
DB	73	ACAACAGAGGAGATGAAGGAGGAGAGATATATCTCTGGAATGGGAGACCCCAACA	132
QY	180	ValPheProAlaPheArgThrAlaValGluAlaGluAspAlaValAlaAlaAlaValArg	199
DB	133	CTCTACTCGTGTTCGTACACACACAGTCTCTTCAAGCTGTTCTGATTCTCTTCTC	192
QY	200	ThrGlyGlnPheAsnGlyProAlaGlyValGlyLeuProAlaAlaArgSerAlaVal	219
DB	193	TCCAAACAAGTCCATGGTGTACTCTCTACCGTCGGTCTTCCCAAGCTCGAAGGCAATA	252
QY	220	AlaGluHisLeuSerGlnGlyValProTyrMetLeuSerAlaAspValPheLeuThr	239
DB	253	GCAGAGTATCTATCCGGTGATCTCCATACAACTTTCACAGGATGATGTTTATCACA	312
QY	240	AlaGlyGlyThrGlnAlaIleGluValIleProValLeuAlaGlnThrAlaGlyAla	259
DB	313	TCGGGTTCCAGCAAGCGATCGATGTAGCATTTGCGATGTAGCTCGTCCACAG--GCT	369
QY	260	AsnIleLeuProArgProGlyTyrProAsnTyrGluAlaArgAlaPheAsnArg	279
DB	370	AATATACTTCTCCAGGCGCTGGTTTCCCAATCTATGAATCTGTCTGTAAGTTTAGACAC	429
QY	280	LeuGluValArgHisPheAspLeuIleProAspLysGlyTyrGluIleAspIleAspSer	299
DB	430	CTTGAAGTTCGTACGTCTCTTCCGGAATAATGGATGGGAGATCGATCTTTGATGCT	489
QY	300	LeuGluSerIleAlaAspLysAsnThrThrAlaMetValIleIleAsnProAsnAsnPro	319
DB	490	GTCGAGGCTCTTCGACAGCAAAACACGGTCTTGTGTTGTTATAAACCCCTGGTAATCCT	549
QY	320	CysGlySerValTyrSerTyrAspHisLeuSerLysValAlaGluValAlaLysArgLeu	339
DB	550	TCCGGGATGTCATAGCTACAGCATTTGTATGAAGATTGCGGATTCGGCGAATAACTA	609
QY	340	GlyIleLeuValIleAlaAspGluValTyrGlyLysLeuValLeuGlySerAlaProPhe	359
DB	610	GGGTTTCTGTGATGTGCTGATGAGGTTTACGGTCATCTTGTCTTGTAGCAAAACCGTTT	669
QY	360	IleProMetGlyValPheGlyHisIleThrProValLeuSerIleGlySerLeuSerLys	379
DB	670	GTGCCAATGGGTGTGTTGGATCTATTGTTCTGCTTACTCTTCTTCTTATCAAG	729
QY	380	SerTyrIleValProGlyTyrArgLeuGlyTyrValAlaValTyrAspProArgLysIle	399
DB	730	AGATGGATAGTTCACAGGTGCGGACTCGGGTGTGTTGTCACCACTGATCTCTCTGTTCC	789
QY	400	LeuGlnGluThrLysIleSerThrSerIleThrAsnTyrLeuAsnValSerThrAspPro	419

Search completed: November 8, 2003, 23:39:00  
Job time: 6141.26 secs

DB	790	TTTAAGGACCCCTAAGATCATTCAGAGGTTTAAAGAAATACTTTTGATATTCTTGGTGGACCA	849
QY	420	AlaThrPheIleGlnAlaAlaLeuProGlnIleLeuGluAsnThrLysGluAspPhePhe	439
DB	850	GCTACATTTATTTCAGGCTGCAGTTCCTCCACTATTTTGAACACAGCGATGAGTCTTCTTC	909
QY	440	LysAlaIleIleGlyLeuLeuLysGluSerSerGluIleCysTyrLysGlnIleLysGlu	459
DB	910	ARGAAAACCTTGAACTCGTTGAAGAACTCTTCGGATATTTGTTGACTTGGATCAAGGAG	969
QY	460	AsnLysTyrIleThrCysProHisLysProGluGlySerMetPheValMetValLysLeu	479
DB	970	ATTCTCTGCAATTGATTCTCTCGCATCCACAGAGGATCCATGGCAATGATGTTAAGCTG	1029
QY	480	AsnLeuHisLeuGluGluIleAspAspIleAspPheCysCysLysLeuAlaLys	499
DB	1030	ATCTCTCATTTACTTGAAGATGTAAGTACGATATCGACTTCTGTTTCAAGTTAGCTAGG	1089
QY	500	GluGluSerValIleLeuCysProGlySerValLeuGlyMetAlaAsnTyrValArgIle	519
DB	1090	GAAGAATCAGTCATCTCTTCTCGGACCGCGGTGGGCTGAAGAACTGGCTGAGGATA	1149
QY	520	ThrPheAlaCysValProSerSerLeuGlnAspGlyLeuGlyArgIleLysSerPheCys	539
DB	1150	ACGTTTGCAGCAGATCGAACTTTCGATTGAAGAAGCTTTTAAAGGATCAAAATGTTTCTAT	1209
QY	540	GlnArgAsnLysLys	544
DB	1210	CTTAGACATGCCAAG	1224